Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

     http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

```
            GenCore version 5.1.6
       Copyright (c) 1993 - 2005 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on: September 22, 2005, 14:39:13 ; Search time 169 Seconds
            (without alignments)
            1375.402 Million cell updates/sec

Title: US-10-712-479-2
Perfect score: 3167
Sequence: 1 MASMSMSYTWTGALITPCAA.........LLSVGVGIYLLPNRHHHHH 601

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_16Dec04:*
            1: geneseqp1980s:*
            2: geneseqp1990s:*
            3: geneseqp2000s:*
            4: geneseqp2001s:*
            5: geneseqp2002s:*
            6: geneseqp2003as:*
            7: geneseqp2003bs:*
            8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3167	100.0	601	8	ADO19021	ADO19021 Full-leng
2	3034.5	95.8	580	8	ADO19023	ADO19023 C-termina
3	3027	95.6	3010	2	AAR30616	AAR30616 Polypepti
4	3016	95.2	591	2	AAW01679	AAW01679 NS5B prot
5	3016	95.2	2201	2	AAW01680	AAW01680 HCV NS2-N
6	3016	95.2	3010	2	AAR20111	AAR20111 Non-A, no
7	3016	95.2	3010	2	AAR20091	AAR20091 Non-A, no
8	3016	95.2	3010	7	ADF88597	ADF88597 Hepatitis
9	3015	95.2	3010	2	AAW98022	AAW98022 Infectiou
10	3015	95.2	3010	4	AAB31170	AAB31170 Amino aci
11	3015	95.2	3010	8	ADO36227	ADO36227 Hepatitis
12	3015	95.2	3010	8	ADO79401	ADO79401 Hepatitis
13	3014	95.2	592	8	ADO36215	ADO36215 Hepatitis
14	3014	95.2	592	8	ADO79389	ADO79389 Hepatitis
15	3014	95.2	3010	2	AAY06423	AAY06423 Non-A, no
16	3013	95.1	3010	2	AAR68864	AAR68864 Hepatitis
17	3012	95.1	591	6	AAB60130	AAB60130 Hepatitis
18	3012	95.1	591	6	AAO26416	AAO26416 Protein d
19	3012	95.1	1985	6	ABU09575	ABU09575 HCV Met-N
20	3012	95.1	3010	2	AAR68622	AAR68622 HCV prote
21	3011	95.1	591	5	AAG79560	AAG79560 HCV NS5B
22	3011	95.1	621	5	AAG79556	AAG79556 HCV decre
23	3011	95.1	621	5	AAE18903	AAE18903 Hepatitis
24	3011	95.1	621	6	ABP71699	ABP71699 HCV NS5B
25	3011	95.1	621	6	ABR41892	ABR41892 Hepatitis
26	3011	95.1	621	6	ABP97967	Abp97967 Amino aci
27	3011	95.1	627	5	AAG79561	Aag79561 HCV decre
28	3011	95.1	627	5	ADG19078	Adg19078 HCV NS5B
29	3011	95.1	3010	2	AAR82694	Aar82694 Partial H
30	3005	94.9	1985	5	AAO18001	Aao18001 Hepatitis
31	3005	94.9	1985	5	AAE15729	Aae15729 Hepatitis
32	3005	94.9	1985	5	AAE15731	Aae15731 Hepatitis
33	3005	94.9	1985	5	AAE15720	Aae15720 Hepatitis
34	3005	94.9	1985	5	AAE15717	Aae15717 Hepatitis
35	3005	94.9	1985	5	AAE15727	Aae15727 Hepatitis
36	3005	94.9	1985	5	AAE15728	Aae15728 Hepatitis
37	3005	94.9	1985	5	AAE15722	Aae15722 Hepatitis
38	3005	94.9	1985	5	AAE15730	Aae15730 Hepatitis
39	3005	94.9	1985	8	ADJ57846	Adj57846 HCV repli
40	3005	94.9	2201	5	ABG30591	Abg30591 Hepatitis
41	3005	94.9	2201	5	ABG30600	Abg30600 Hepatitis
42	3005	94.9	2201	5	ABG30581	Abg30581 Hepatitis
43	3005	94.9	2201	5	ABG30586	Abg30586 Hepatitis
44	3005	94.9	2201	5	ABG30593	Abg30593 Hepatitis
45	3005	94.9	2201	5	ABG30582	Abg30582 Hepatitis

ALIGNMENTS

RESULT 1
ADO19021
ID   ADO19021 standard; protein; 601 AA.
AC   ADO19021;
XX
DT   12-AUG-2004 (first entry)
XX
DE   Full-length HCV NS5B polymerase, FL NS5B.
XX
KW   RNA polymerase activity; continuous-read assay;
KW   Hepatitis C virus infection; HCV infection; polymerase; NS5B;
KW   bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
KW   diabetes; ocular disorder; renal dysfunction; lymphoma;
KW   lymphoproliferative disorder; metabolic disorder; arthritis;
KW   sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
KW   virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;
KW   nephrotropic; antiarthritic; enzyme.
XX
OS   Hepatitis C virus.
XX
PN   WO200404228-A2.
XX
PD   27-MAY-2004.
XX
PF   13-NOV-2003; 2003WO-US036465.
XX
PR   13-NOV-2002; 2002US-0425981P.
XX
PA   (PFIZ) PFIZER INC.
XX
PI   Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;
PI   Morgan AG;
XX
DR   WPI; 2004-420337/39.
DR   N-PSDB; ADO19020.
XX
PT   Detecting RNA polymerase activity in a continuous-read manner, useful in
PT   treating osteoporosis, carcinomas, cardiovascular diseases, ocular
PT   disorders or arthritis, by contacting an RNA polymerase with an
PT   oligonucleotide template.
XX
PS   Claim 5; SEQ ID NO 2; 46pp; English.
XX
CC   The present invention relates to a method for detecting RNA polymerase
CC   activity in a continuous-read manner. The method comprises contacting an
CC   RNA polymerase with an oligonucleotide template in a reaction mixture
CC   comprising an assay buffer, under conditions in which the RNA polymerase

```
CC   is active, adding a fluorescent dye capable of binding double-stranded
CC   nucleic acid molecules to the reaction mixture, and measuring the
CC   fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
CC   C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
CC   infection, bone mineral diseases like osteoporosis, carcinomas,
CC   cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
CC   lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
CC   sleep disorders and thyroid disorders. The present sequence represents
CC   full-length HCV NS5B polymerase (designated FL NS5B).
XX
SQ   Sequence 601 AA;

Query Match             100.0%;  Score 3167;  DB 8;  Length 601;
Best Local Similarity   100.0%;  Pred. No. 4e-294;
Matches   601;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60

Qy        61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120

Qy       121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180

Qy       181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240

Qy       241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300

Qy       301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360

Qy       361 EYDLELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 EYDLELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420

Qy       421 PTLWARMILMWHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 PTLWARMILMWHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480

Qy       481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICCKYLFNWAVRTKLKL 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICCKYLFNWAVRTKLKL 540

Qy       541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNRHHHHH 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNRHHHHH 600

Qy       601 H 601
             |
Db       601 H 601
```

RESULT 2
AD019023
ID   AD019023 standard; protein; 580 AA.
XX
AC   AD019023;
XX
DT   12-AUG-2004 (first entry)
XX
DE   C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.
XX
KW   RNA polymerase activity; continuous-read assay;
KW   Hepatitis C virus infection; HCV infection; polymerase; NS5B;

```
KW   bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
KW   diabetes; ocular disorder; renal dysfunction; lymphoma;
KW   lymphoproliferative disorder; metabolic disorder; arthritis;
KW   sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
KW   virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;
KW   nephrotropic; antiarthritic; enzyme.
XX
OS   Hepatitis C virus.
XX
PN   WO200404228-A2.
XX
PD   27-MAY-2004.
XX
PF   13-NOV-2003; 2003WO-US036465.
XX
PR   13-NOV-2002; 2002US-0425981P.
XX
PA   (PFIZ) PFIZER INC.
XX
PI   Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;
PI   Morgan AG;
XX
DR   WPI; 2004-420337/39.
DR   N-PSDB; AD019022.
XX
PT   Detecting RNA polymerase activity in a continuous-read manner, useful in
PT   treating osteoporosis, carcinomas, cardiovascular diseases, ocular
PT   disorders or arthritis, by contacting an RNA polymerase with an
PT   oligonucleotide template.
XX
PS   Claim 7; SEQ ID NO 4; 46pp; English.
XX
CC   The present invention relates to a method for detecting RNA polymerase
CC   activity in a continuous-read manner. The method comprises contacting an
CC   RNA polymerase with an oligonucleotide template in a reaction mixture
CC   comprising an assay buffer, under conditions in which the RNA polymerase
CC   is active, adding a fluorescent dye capable of binding double-stranded
CC   nucleic acid molecules to the reaction mixture, and measuring the
CC   fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
CC   C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
CC   infection, bone mineral diseases like osteoporosis, carcinomas,
CC   cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
CC   lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
CC   sleep disorders and thyroid disorders. The present sequence represents C-
CC   terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).
XX
SQ   Sequence 580 AA;

Query Match             95.8%;  Score 3034.5;  DB 8;  Length 580;
Best Local Similarity   96.5%;  Pred. No. 1.9e-281;
Matches   580;  Conservative   0;  Mismatches   0;  Indels   21;  Gaps   1;

Qy         1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60

Qy        61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120

Qy       121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180

Qy       181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240

Qy       241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300

Qy       301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
```

SQ   Sequence 3010 AA;

Query Match   95.6%; Score 3027; DB 2; Length 3010;
Best Local Similarity   96.8%; Pred. No. 1.3e-279;
Matches   572; Conservative   9; Mismatches   10; Indels   0; Gaps   0;

```
Qy       5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           |||||||||||||||||||||||| :|||||||||| :| :||||| ||||||||||||||
Db    2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVATTSRSAGLRQKKVTFDRLQVL 2479

Qy      65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVTHIR 2539

Qy     125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||| || | ||||||||||||||||||||||||||||||||||||||||||||
Db    2540 SVWKDLLEDTETPISTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy     185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy     245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||||:  ||||||||:|||| ||||||||||||||||||||||||||||||||| :
Db    2660 QCCDLAPEAKLAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy     305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2720 AACRAAKLRDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy     365 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2840 ARMILMTHFFSILLAQEQLEKTLDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||  ||||||||||||||||||| ||||||||||||||||||||
Db    2900 EINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 4
AAW01679
ID   AAW01679 standard; protein; 591 AA.
XX
AC   AAW01679;
XX
DT   17-OCT-2003 (revised)
DT   25-MAR-2003 (revised)
DT   19-AUG-1997 (first entry)
XX
DE   NS5B protein (residues 2420-3010 of the HCV polyprotein).
XX
KW   Hepatitis C virus; HCV; non-structural protein; NS5B; polyprotein;
KW   RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;
KW   TNTase; method; assay; in vitro activity; therapy; inhibitor.
XX
OS   Hepatitis C virus; strain BK.
XX
PN   WO9637619-A1.
XX
PD   28-NOV-1996.
XX
PF   24-MAY-1996; 96WO-IT000106.
XX
PR   25-MAY-1995; 95IT-RM000343.
XX

```
Db     301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYA 420

Qy     421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480

Qy     481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540

Qy     541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNRHHHHH 600
           |||||||||||||||||||||||||||||||||                    ||||||
Db     541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARP------------------RHHHHH 579

Qy     601 H 601
             |
Db     580 H 580
```

RESULT 3
AAR30616
ID   AAR30616 standard; protein; 3010 AA.
XX
AC   AAR30616;
XX
DT   27-AUG-2003 (revised)
DT   25-MAR-2003 (revised)
DT   19-MAY-1993 (first entry)
XX
DE   Polypeptide coded by Korean HCV full cDNA sequence LBC1.
XX
KW   KHCV-LBC1; diagnosis; vaccine.
XX
OS   Hepatitis C virus.
XX
PN   EP521318-A2.
XX
PD   07-JAN-1993.
XX
PF   10-JUN-1992; 92EP-00109753.
XX
PR   10-JUN-1991; 91KR-00009510.
PR   06-AUG-1991; 91KR-00013601.
XX
PA   (LUCK-) LUCKY LTD.
XX
PI   Cho JM, Lee YB, Park YW, Lim KJ, Choi DY, So HS, Kim CH;
PI   Kim ST, Yang JY;
XX
DR   WPI; 1993-001883/01.
DR   N-PSDB; AAQ33282.
XX
PT   DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV) - for
PT   diagnosing and vaccinating against KHCV infections.
XX
PS   Disclosure; Fig 2; 119pp; English.
XX
CC   The polypeptide is that encoded by the full cDNA sequence of Korean
CC   hepatitis C virus (KHCV) cDNA, KHCV-LBC1. It or its fragments may be used
CC   in a specific and accurate method for detecting KCHV antibodies in the
CC   serum of hepatitis C patients. Antibodies directed against these
CC   polypeptides are useful for the purification of KCHV antigens and for the
CC   development of an improved diagnostic to detect KHCV antigens in a
CC   sample. The polypeptides may also be used in a vaccine for treatment and
CC   prevention of KHCV infection at a dosage of 5-200 ug/peptide. (Updated on
CC   25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
CC   field.)
XX

PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Francesco R, Tomei L, Behrens S;
XX
DR   WPI; 1997-021225/02.
XX
PT   Reproducing enzymatic activities of HCV in vitro - using sequences contg.
PT   NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl
PT   transferase activities.
XX
PS   Claim 1; Page 24-26; 49pp; English.
XX
CC   A novel method for reproducing in vitro the RNA-dependent RNA polymerase
CC   (RdRp) activity or the terminal nucleotidyl transferase (TNTase) activity
CC   encoded by hepatitis C virus (HCV), is characterised in that sequences
CC   contg. NS5B are used in the mixture. The method is used for assaying in
CC   vitro the activities of RdRp and TNTase encoded by HCV in order to
CC   identify, for therapeutic purposes, compounds that inhibit these
CC   enzymatic activities and therefore might interfere with the replication
CC   of the HCV. The present sequence comprises amino acids 2420-3010 of the
CC   HCV polyprotein and corresponds to the NS5B protein. cDNA encoding this
CC   protein was cloned between the BamHI and HindIII sites of pBlueBacIII to
CC   form pBac5B. Another expression plasmid, pBac25 (containing cDNA encoding
CC   amino acids 810-3010 of HCV; see AAW01680) was also constructed. Extracts
CC   of Bac25- or Bac5B-infected Sf9 cells contain a novel magnesium-dependent
CC   enzymatic activity that catalyses de novo RNA synthesis. This activity
CC   was shown to be dependent on the presence of RNA, but independent of an
CC   added primer or of the origin of the input RNA molecule. As the products
CC   generated by extracts of Sf9 cells infected with either Bac25 or Bac5
CC   appeared to be identical, the experiments indicated that the observed
CC   RdRp activity is encoded by the HCV NS5B protein. (Updated on 25-MAR-2003
CC   to correct PR field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 591 AA;

Query Match   95.2%; Score 3016; DB 2; Length 591;
Best Local Similarity 96.6%; Pred. No. 1.2e-279;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          ||||||||||||||||||||||||||||||||:||:|||||||| :||||||||||||
Db   1 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 60

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          ||||||||||||||||||||||||||||||||||||||||||||||:|||:::||||||
Db   61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 120

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy 185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||:||:||||||||||||:||||||||||:::||||||||||||||||||||||||||||
Db 181 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 240

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||:|||||||:||||:||||||||||||||||||||||||||||||||:
Db 241 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 360

Qy 365 ELITCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db 361 ELITCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 420

Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy 485 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||||||||||||:|||||||||||:||||||||
Db 481 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 540

Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 591

RESULT 5
AAW01680
ID   AAW01680 standard; protein; 2201 AA.
XX
AC   AAW01680;
XX
DT   17-OCT-2003 (revised)
DT   25-MAR-2003 (revised)
DT   19-AUG-1997 (first entry)
XX
DE   HCV NS2-NS5B non-structural protein.
XX
KW   Hepatitis C virus; HCV; non-structural protein; NS5B; polyprotein;
KW   RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;
KW   TNTase; method; assay; in vitro activity; therapy; inhibitor.
XX
OS   Hepatitis C virus; strain BK.
XX
PN   WO9637619-A1.
XX
PD   28-NOV-1996.
XX
PF   24-MAY-1996; 96WO-IT000106.
XX
PR   25-MAY-1995; 95IT-RM000343.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Francesco R, Tomei L, Behrens S;
XX
DR   WPI; 1997-021225/02.
XX
PT   Reproducing enzymatic activities of HCV in vitro - using sequences contg.
PT   NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl
PT   transferase activities.
XX
PS   Example 1; Page 26-34; 49pp; English.
XX
CC   A novel method for reproducing in vitro the RNA-dependent RNA polymerase
CC   (RdRp) activity or the terminal nucleotidyl transferase (TNTase) activity
CC   encoded by hepatitis C virus (HCV), is characterised in that sequences
CC   contg. NS5B are used in the mixture. The method is used for assaying in
CC   vitro the activities of RdRp and TNTase encoded by HCV in order to
CC   identify, for therapeutic purposes, compounds that inhibit these
CC   enzymatic activities and therefore might interfere with the replication
CC   of the HCV. The present sequence comprises amino acids 810-3010 of the
CC   HCV polyprotein and corresponds to NS2-NS5B proteins. cDNA encoding this
CC   protein was cloned between the NcoI and HindIII sites of pBlueBacIII to
CC   form pBac25. Another expression plasmid, pBac5B (containing cDNA encoding
CC   amino acids 2420-3010 of HCV; see AAW01679) was also constructed.
CC   Extracts of Bac25- or Bac5B-infected Sf9 cells contain a novel magnesium-
CC   dependent enzymatic activity that catalyses de novo RNA synthesis. This
CC   activity was shown to be dependent on the presence of RNA, but
CC   independent of an added primer or of the origin of the input RNA
CC   molecule. As the products generated by extracts of Sf9 cells infected
CC   with either Bac25 or Bac5 appeared to be identical, the experiments
CC   indicated that the observed RdRp activity is encoded by the HCV NS5B
CC   protein. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT
CC   -2003 to standardise OS field)
XX
SQ   Sequence 2201 AA;

Query Match   95.2%; Score 3016; DB 2; Length 2201;
Best Local Similarity 96.6%; Pred. No. 9.1e-279;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

```
Qy      5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
              ||||||||||||||||||||||||||||||||||||| ::|||||||| |||||||||||
Db   1611  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMYATTSRSAGLRQKKVTFDRLQVL  1670

Qy     65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
              ||||||||||||||||||||||||||||||||||||||||||||||||| ::|::|||||:
Db   1671  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH  1730

Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
              |||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||
Db   1731  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  1790

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
              |||| :|||||||||||:||||||||||:||:||||||||||||||||||||||||||||
Db   1791  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY   1850

Qy    245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
              |||||||||||||:||||||||:|||:|||||||||||||||||||||||||||||||||:
Db   1851  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  1910

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  1970

Qy    365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
              |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   1971  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  2030

Qy    425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  2090

Qy    485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
              |||||||||||||||||||||||||||||:|||||||||:||||||||||:|||||||||
Db   2091  EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP  2150

Qy    545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  2201
```

RESULT 6
AAR20111

ID   AAR20111 standard; protein; 3010 AA.
XX
AC   AAR20111;
XX
DT   25-MAR-2003 (revised)
DT   01-MAY-1992 (first entry)
XX
DE   Non-A, non-B viral genome product.
XX
KW   NANBV; vaccine; immunodiagnosis; antigen; antibody.
XX
OS   Non-A,
OS   non-B hepatitis virus.
XX
FH   Key   Location/Qualifiers
FH
FT   Protein   1..115
FT   /label= C
FT   /note= "core protein"
FT   Protein   116..191
FT   /label= M
FT   /note= "matrix protein"
FT   Protein   192..389
FT   /label= E
FT   /note= "envelope protein"
FT   Protein   390..729
FT   /label= NS1
FT   Protein   730..1006
FT   /label= NS2
FT   Protein   1007..1614
FT   /label= NS3
FT   Protein   1615..1862
FT   /label= NS4a
FT   Protein   1863..2012
FT   /label= NS4b
FT   Protein   2013..3010
FT   /label= NS5
XX
PN   EP463848-A.
XX
PD   02-JAN-1992.
XX
PF   25-JUN-1991;   91EP-00305717.
XX
PR   25-JUN-1990;
PR   31-AUG-1990;   90JP-00167466.
PR   09-NOV-1990;   90JP-00230921.
PR   28-DEC-1990;   90JP-00305605.
PR   08-MAY-1991;   90US-00635451.
PR   14-MAY-1991;   91JP-00132090.
XX   91JP-00138493.
PA   (OSAU) UNIV OSAKA.
PA   (REMI-) RES FOUND MICROBIAL.
XX
DR   WPI; 1992-009412/02.
DR   N-PSDB; AAQ20268.
XX
PT   Non-A, non-B hepatitis virus (NANBV) particles - as vaccines, immuno-
PT   diagnostics and screening agents for NANBV, and to remove NANBV from
PT   blood.
XX
PS   Disclosure; Fig 2; 89pp; English.
XX
CC   The sequence (SEQ ID NO 1) was deduced from several overlapping from a
CC   library prepd. from NANBV RNA. Antigenic polypeptides from the sequence
CC   can be used as immunoassay reagents, for screening donated blood, and as
CC   immunogens for vaccine prodn. Antibodies raised to the peptides can be
CC   used in immunoassays to detect or quantify NANBV antigens in liver tissue
CC   and blood. Preferred poly- peptides include residues 1-30, -115, or 2012;
CC   47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
CC   729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC   and 2013-3010. The sequence is also disclosed in EP-464287 (SEQ ID NO 1).
CC   See AAR20091 for details of this specification. (Updated on 25-MAR-2003
CC   to correct PD field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 3010 AA;

Query Match   95.2%; Score 3016; DB 2; Length 3010;
Best Local Similarity   96.6%; Pred. No. 1.5e-278;
Matches   571; Conservative   11; Mismatches   9; Indels   0; Gaps   0;

```
Qy      5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
              ||||||||||||||||||||||||||||||||||||| ::|||||||| |||||||||||
Db   2420  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMYATTSRSAGLRQKKVTFDRLQVL  2479

Qy     65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
              ||||||||||||||||||||||||||||||||||||||||||||||||| ::|::|||||:
Db   2480  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH  2539

Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
              |||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   2540  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  2599

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
              |||| :|||||||||||:||||||||||:||:||||||||||||||||||||||||||||
Db   2600  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY   2659

Qy    245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
              |||||||||||||:||||||||:|||:|||||||||||||||||||||||||||||||||:
Db   2660  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  2719

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  2779
```

PT   New DNA from non-A, non-B hepatitis virus - and derived antigenic
PT   polypeptide(s) useful for diagnostics, blood screening and in vaccines.
XX
PS   Claim 3; Fig 2; 89pp; English.
XX
CC   The sequence was deduced from several overlapping "BK" cDNA clones obtd.
CC   by "gene walking" using a cDNA clone isolated from a library prepd. from
CC   NANBV RNA. Antigenic polypeptides from the sequence can be used as
CC   immunoassay reagents, for screening donated blood, and as immunogens for
CC   vaccine prodn. Antibodies raised to the peptides can be used in
CC   immunoassays to detect or quantify NANBV antigens in liver tissue and
CC   blood. Preferred polypeptides are include residues 1-30, -115, or 2012;
CC   47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
CC   729; 730-1005; 1006- 1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC   and 2013-3010. The sequence is also disclosed in EP-463848 (SEQ ID NO 1)
CC   in which a virus particle contg. antigens encoded by the sequence is
CC   claimed. See AAR20011 for details of this specification. (Updated on 25-
CC   MAR-2003 to correct PA field.)
XX
SQ   Sequence 3010 AA;

Query Match   95.2%; Score 3016; DB 2; Length 3010;
Best Local Similarity   96.6%; Pred. No. 1.5e-278;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           |||||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVS 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTVPIDTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCERMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||||||||:||||||||||||||||:||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010

RESULT 8
ADF88597
ID   ADF88597 standard; protein; 3010 AA.
XX
AC   ADF88597;
XX
DT   26-FEB-2004 (first entry)

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||||||||:||||||||||||||||:||||||||||:|||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010

RESULT 7
AAR20091
ID   AAR20091 standard; protein; 3010 AA.
XX
AC   AAR20091;
XX
DT   25-MAR-2003 (revised)
DT   01-MAY-1992 (first entry)
XX
DE   Non-A, non-B viral genome product.
XX
KW   NANBV; vaccine; immunodiagnosis; antigen; antibody.
XX
OS   Non-A.
OS   non-B hepatitis virus.
XX
FH   Key   Location/Qualifiers
FT   Protein   1. .115
FT   /label= C
FT   /note= "core protein"
FT   Protein   116. .191
FT   /label= M
FT   /note= "matrix protein"
FT   Protein   192. .389
FT   /label= E
FT   /note= "envelope protein"
FT   Protein   390. .729
FT   /label= NS1
FT   Protein   730. .1006
FT   /label= NS2
FT   Protein   1007. .1614
FT   /label= NS3
FT   Protein   1615. .1862
FT   /label= NS4a
FT   Protein   1863. .2012
FT   /label= NS4b
FT   Protein   2013. .3010
FT   /label= NS5
XX
PN   EP464287-A.
XX
PD   08-JAN-1992.
XX
PF   28-DEC-1990; 90EP-00314371.
XX
PR   25-JUN-1990; 90JP-00167466.
PR   31-AUG-1990; 90JP-00230921.
PR   09-NOV-1990; 90JP-00305605.
PR   17-JUN-1991; 91EP-00401604.
XX
PA   (OSAU) UNIV OSAKA.
XX
DR   WPI; 1992-009617/02.
DR   N-PSDB; AAQ21829.
XX

```
XX
DE   Hepatitis C virus NS3 gene protein, SEQ ID No 6.
XX
KW   liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.
XX
OS   Hepatitis C virus.
XX
PN   JP200321018l-A.
XX
PD   29-JUL-2003.
XX
PF   30-MAY-2002; 2002JP-00158335.
XX
PR   16-NOV-2001; 2001JP-00352443.
XX
PA   (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX
DR   WPI; 2003-819836/77.
DR   N-PSDB; ADF88596.
XX
PT   Diagnosing liver cancer, involves amplifying amino terminal region of
PT   hepatitis-C virus gene using predetermined primer and determining
PT   hepatitis-C virus in a base sequence of obtained DNA fragment.
XX
PS   Disclosure; SEQ ID NO 6; 36pp; Japanese.
XX
CC   The invention relates to the novel testing method for diagnosing liver
CC   cancer. The novel method comprises amplifying the amino terminal region
CC   of a hepatitis-C virus NS3 gene using a predetermined primer and
CC   determining the hepatitis-C virus in a base sequence of the obtained DNA
CC   fragment. The novel testing method is useful for diagnosing liver cancer
CC   and also used in a gene amplification technique, a clinical laboratory
CC   test reagent, a polymerase chain reaction, a base sequence analysis and
CC   genetic engineering. The method enables the detection of a hepatitis-C
CC   virus having high carcinogenicity with high specificity. This sequence
CC   represents the protein of the hepatitis-C virus NS3 gene of the
CC   invention.
XX
SQ   Sequence 3010 AA;

Query Match             95.2%;  Score 3016;  DB 7;  Length 3010;
Best Local Similarity   96.3%;  Pred. No. 1.5e-278;
Matches   569;  Conservative   12;  Mismatches   10;  Indels   0;  Gaps   0;

Qy        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMVYSTTSRSASLRQKKVTFDRLQVL 2479

Qy       65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR  124
            ||||||||||||||||||||:||||:||||||||||||||||||||||:|||||||||||||
Db     2480 DDHYRDVLKEMKAKASTVKAKLSIEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR 2539

Qy      125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
            ||||:||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2540 SVWEDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy      185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEBESIY  244
            ||||||||||||||||:||:|||||||||:||:||||||||||||||||||||||||||||
Db     2600 TLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKCCMGFSYDTRCFDSIVTENDIRTEESIY 2659

Qy      245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
            ||||||||||||||||||||||:||||||||||||||||||||||||||||:|||||||:|
Db     2660 QCCDLAPEARQAIRSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy      305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
            |||||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||||
Db     2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy      365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETVRHTPVNSWLGNIIMYAPTLW 2839

Qy      425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
                |||||||||||||||||||||||||||||||||||||||| :||:||||||
Db     2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy      545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
                ||:|||||||||||:||||||||||||||||||||||:||||||||||||
Db     2960 AASQLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010


RESULT 9
AAW98022
ID   AAW98022  standard; protein; 3010 AA.
XX
AC   AAW98022;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Infectious hepatitis C virus genotype 1b strain HC-J4 protein.
XX
KW   HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
KW   assay; antiviral; virucide.
XX
OS   Hepatitis C virus.
XX
XX
PN   WO9904008-A2.
XX
PD   28-JAN-1999.
XX
PF   16-JUL-1998;    98WO-US014688.
XX
PR   18-JUL-1997;    97US-0053062P.
PR   27-JAN-1998;    98US-0001441G.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
DR   WPI; 1999-132252/11.
DR   N-PSDB; AAX24843.
XX
PT   New isolated hepatitis C virus nucleic acids - used to develop products
PT   for the diagnosis, prevention and treatment of HCV infections and for
PT   developing screening assays.
XX
PS   Claim 2; Fig 14G-H; 126pp; English.
XX
CC   This protein is encoded by the infectious hepatitis C virus (HCV)
CC   genotype 1b strain HC-J4 genome (see AAX24833). HC-J4 was obtained from
CC   acute phase plasma of a chimpanzee infected with serum containing HC-
CC   J4/91. The infectious nucleic acid sequence can be used to produce
CC   chimeric genomes (see AAX24833) consisting of the open reading frames of
CC   infectious nucleic acid sequences of other genotypes (including genotypes
CC   1-6) and subtypes (such as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV.
CC   The invention also relates to the introduction of mutations or deletions
CC   into infectious nucleic acid sequences in order to produce an attenuated
CC   HCV virus suitable for vaccine development. Infectious nucleic acid
CC   sequences can also be used to produce attenuated virus via passage in
CC   vitro or in vivo of the viruses produced by transfection of a host cell
CC   with the infectious nucleic acid sequence. Vaccines comprising one or
CC   more polypeptides made from the infectious nucleic acid sequence are used
CC   to immunise mammals, especially humans, against hepatitis C. The nucleic
CC   acid sequences can also be used to induce protective immunity against the
CC   virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
CC   protease) can additionally be used to develop screening assays to
CC   identify antiviral agents for HCV
XX
SQ   Sequence 3010 AA;

Query Match           95.2%;     Score 3015;   DB 2;  Length 3010;
Best Local Similarity 95.9%;     Pred. No. 1.9e-278;
```

PT   New nucleic acid comprising a chimeric bovine viral diarrhea virus genome
PT   in which the (non-)structural region has been replaced by hepatitis C
PT   virus (HCV) genome useful for treating or preventing HCV signs and
PT   symptoms.
XX
PS   Disclosure; Fig 4G-H; 97pp; English.
XX
CC   The specification describes a nucleic acid comprising a chimeric virus
CC   genome, specifically bovine viral diarrhea virus (BVDV) genome in which
CC   the (non-)structural region has been replaced by the (non-)structural
CC   region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC   the chimeric virus and the chimeric virus are useful for identifying cell
CC   lines capable of supporting the replication of these chimeric viruses, in
CC   screening for neutralizing antibodies to HCV of different genotypes, in
CC   the production of HCV-BVDV virions, for the development of inactivated or
CC   attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
CC   molecular properties of HCV indirectly in vitro, and in identifying
CC   inhibitors of viral enzyme activity which would be useful as antiviral
CC   agents. Formulations or compositions comprising the chimeric virions may
CC   be used to treat or prevent the signs and symptoms of HCV. The present
CC   sequence is encoded by a HCV clone, which is used to construct chimeric
CC   nucleic acids of the invention
XX
SQ   Sequence 3010 AA;

Query Match   95.2%;   Score 3015;   DB 4;   Length 3010;
Best Local Similarity   95.9%;   Pred. No. 1.9e-278;
Matches   567;   Conservative   16;   Mismatches   8;   Indels   0;   Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
       ||||||||||||||| :||||||||||||| ||||||||||| |||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
       |||||||||||||| ||||||||||| |||||||||||||||||||| ||||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIR 2539

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
       ||| |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
       |||||||||||||||||||||||||||| :|||||||||||||||||:|:||||||||||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
       ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
       |||||||||||||||||||||||||||||:::||| ||| :|||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
       |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
       ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFTLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
       |||||||||||||||||| |||||||||||||||||||||||| |||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
       |||:||||||||||||||||||||||||||||| ||||||||||||||||
Db   2960 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLLSVGVGIYLLPNR 3010

RESULT 11
ADO36227
ID   ADO36227 standard; protein; 3010 AA.

RESULT 10
AAB31170
ID   AAB31170 standard; protein; 3010 AA.
XX
AC   AAB31170;
XX
DT   02-APR-2001 (first entry)
XX
DE   Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1b.
XX
KW   Chimeric virus; bovine viral diarrhea virus; BVDV; hepatitis C virus;
KW   HCV; vaccine; viral inhibitor; antiviral.
XX
OS   Hepatitis C virus.
XX
PN   WO200075352-A2.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015527.
XX
PR   04-JUN-1999; 99US-0137817P.
XX
PA   (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nam J, Bukh J, Emerson SU, Purcell RH;
XX
DR   WPI; 2001-071081/08.
DR   N-PSDB; AAC86939.
XX

```
XX
AC   AD036227;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Hepatitis C virus (HCV) J4L6 wild-type polyprotein.
XX
KW   hepatotropic; virucide; vaccine; gene therapy; vaccine;
KW   Hepatitis C virus; HCV; core protein; HCV infection; vaccination;
KW   polyprotein.
XX
OS   Hepatitis C virus.
XX
PN   WO2004046175-A1.
XX
PD   03-JUN-2004.
XX
PF   13-NOV-2003; 2003WO-EP012793.
XX
PR   15-NOV-2002; 2002GB-00026722.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Brett S,   Hamblin PA,   Ogilvie L;
XX
DR   WPI; 2004-420613/39.
DR   N-PSDB; AD036222.
XX
PT   New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT   the polypeptide sequences of the HCV core and at least one other HCV
PT   protein, for use in medicine, particularly for manufacturing a medicament
PT   for treating HCV.
XX
PS   Disclosure; Fig 6; 78pp; English.
XX
CC   The invention describes a polynucleotide vaccine comprising a
CC   polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
CC   protein and at least 1 other HCV protein, and causes expression of the
CC   proteins in cells (in which (S1) has been mutated or positioned relative
CC   to the polynucleotide sequence encoding the other HCV protein, so that
CC   the negative effect of the Core protein on expression of the other HCV
CC   protein is reduced). Also described are: a method of preventing or
CC   treating an HCV infection in a mammal, comprising administering the
CC   vaccine cited above to a mammal; and a method of vaccination of an
CC   individual, comprising taking a polynucleotide vaccine as cited above,
CC   coating the polynucleotide onto gold beads and delivering the gold beads
CC   into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
CC   antibodies used in the methods, are also disclosed. The polynucleotide
CC   vaccine is useful in the manufacture of a medicament for the treatment of
CC   HCV. This is the amino acid sequence of the wild type HCV polyprotein.
XX
SQ   Sequence 3010 AA;

Query Match             95.2%;  Score 3015;  DB 8;  Length 3010;
Best Local Similarity   95.9%;  Pred. No. 1.9e-278;
Matches  567;  Conservative   16;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          ||||||||||||||||||||||||  ||||||||||  || | |||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVBEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          ||||||||||||||||||||||||  |||||||||||||||||||| |||| ||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIR 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          || |||||| ||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSVDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||||||||||||||| ||||||||||||||||||| ||||||||||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 2659
```

```
Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||||||||||||| |||||||||||||||| | ||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWIGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFTLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||| ||| ||||||||||||||| ||||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||  |||||||||||||||||||||||||||| | |||||||||||||||||
Db   2960 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFPICLLLLSVGVGIYLLPNR 3010
```

```
RESULT 12
AD079401
ID   AD079401 standard; protein; 3010 AA.
XX
AC   AD079401;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Hepatitis C virus J4L6 genome wild-type polyprotein.
XX
KW   HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
PN   WO2004046176-A1.
XX
PD   03-JUN-2004.
XX
PF   13-NOV-2003; 2003WO-EP012830.
XX
PR   15-NOV-2002; 2002GB-00026722.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Brett S,   Hamblin PA,   Ogilvie L;
XX
DR   WPI; 2004-420614/39.
DR   N-PSDB; AD079396.
XX
PT   New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT   the polypeptide sequences of the HCV proteins Core, NS3, NS4B and NS5B,
PT   for use in medicine, in particular for manufacturing a medicament for the
PT   treatment of HCV.
XX
PS   Disclosure; Fig 6; 79pp; English.
XX
CC   The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type
CC   polyprotein sequence. HCV DNA vaccines of the invention comprise a
CC   polynucleotide that encodes the HCV proteins Core, NS3, NS4B and NS5B,
CC   and does not encode the NS4A and/or NS5A proteins. Preferably, the codon
CC   usage of the polynucleotide sequence resembles that of highly expressed
CC   human genes. The polynucleotides may encode individual proteins or fusion
CC   proteins. Preferred fusions include double fusions between NS4B and NS5B
CC   and between Core and NS3. The vaccines are useful for the treatment or
CC   prevention of an HCV infection.
XX
SQ   Sequence 3010 AA;
```

DR   WPI; 2004-420613/39.
DR   N-PSDB; AD036226.
XX
PT   New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT   the polypeptide sequences of the HCV core and at least one other HCV
PT   protein, for use in medicine, particularly for manufacturing a medicament
PT   for treating HCV.
XX
PS   Example 2; Page 27; 78pp; English.
XX
CC   The invention describes a polynucleotide vaccine comprising a
CC   polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
CC   protein and at least 1 other HCV protein, and causes expression of the
CC   proteins in cells (in which (S1) has been mutated or positioned relative
CC   to the polynucleotide sequence encoding the other HCV protein, so that
CC   the negative effect of the Core protein on expression of the other HCV
CC   protein is reduced) .Also described are: a method of preventing or
CC   treating an HCV infection in a mammal, comprising administering the
CC   vaccine cited above to a mammal; and a method of vaccination of an
CC   individual, comprising taking a polynucleotide vaccine as cited above,
CC   coating the polynucleotide onto gold beads and delivering the gold beads
CC   into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
CC   antibodies used in the methods, are also disclosed. The polynucleotide
CC   vaccine is useful in the manufacture of a medicament for the treatment of
CC   HCV. This is the amino acid sequence of the HCV NSSB protein.
XX
SQ   Sequence 592 AA;
XX

Query Match   95.2%; Score 3014; DB 8; Length 592;
Best Local Similarity   96.1%; Pred. No. 1.8e-279;
Matches 569; Conservative 14; Mismatches   9; Indels   0; Gaps   0;

Qy   4 MSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQV 63
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1 MSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQV 60

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHI 123
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   61 LDDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGARDVRNLSSRAVNHI 120

Qy   124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 183
          |||| ||||| ||| ||||||||:||||||||||||||||||||||||||||||||||||
Db   121 RSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 180

Qy   184 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
          |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db   181 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFGSTVTESDIRVEESI 240

Qy   244 YQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 YQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 300

Qy   304 AAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYD 363
          |:|||||||||||||||||||||||||||||||:|||:||||||||||||||||||||||
Db   301 TAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYD 360

Qy   364 LELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTL 423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 LELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTL 420

Qy   424 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 483
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 480

Qy   484 GEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPI 543
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   481 GEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPI 540

Qy   544 PAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||:|||||||||||||||||||||||||||| |||||||||||||||||
Db   541 PAASQLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLLSVGVGIYLLPNR 592

Query Match   95.2%; Score 3015; DB 8; Length 3010;
Best Local Similarity   95.9%; Pred. No. 1.9e-278;
Matches 567; Conservative 16; Mismatches   8; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||| ||||||||| |||||||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGARDVRNLSSRAVNHIR 2539

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||| ||||| ||| |||||||:|||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||||||||||||||||||||||||||||||:|||||||||||||| ||||||||||||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFGSTVTESDIRVEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||:||||:||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFTLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||:|||||||||||||||||||||||||||| |||||||||||||||||||
Db   2960 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLLSVGVGIYLLPNR 3010

RESULT 13
AD036215
ID   AD036215 standard; protein; 592 AA.
XX
AC   AD036215;
XX
DT   26-AUG-2004 (first entry)
XX
DE   Hepatitis C virus (HCV) NS5B protein.
XX
KW   hepatotropic; virucide; vaccine; gene therapy; vaccine;
KW   Hepatitis C virus; HCV; core protein; HCV infection; vaccination; NS5B.
XX
OS   Hepatitis C virus.
XX
PN   WO200404175-A1.
XX
PD   03-JUN-2004.
XX
PF   13-NOV-2003; 2003WO-EP012793.
XX
PR   15-NOV-2002; 2002GB-00026722.
XX
PA   (GLAX) GLAXO GROUP LTD.
XX
PI   Brett S, Hamblin PA, Ogilvie L;
XX

RESULT 14
ADO79389
ID   ADO79389 standard; protein; 592 AA.
XX
AC   ADO79389;
XX
DT   26-AUG-2004 (first entry)
XX
DE   Hepatitis C virus NS5B protein.
XX
KW   HCV; NS5B; vaccine; DNA immunisation; hepatotropic; virucide; mutant;
KW   mutein.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
PN   WO200404176-A1.
XX
PD   03-JUN-2004.
XX
PF   13-NOV-2003; 2003WO-EP012830.
XX
PR   15-NOV-2002; 2002GB-00026722.
XX
PA   (GLAX) GLAXO GROUP LTD.
XX
PI   Brett S, Hamblin PA, Ogilvie L;
XX
DR   WPI; 2004-420614/39.
DR   N-PSDB; ADO79400.
XX
PT   New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT   the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
PT   for use in medicine, in particular for manufacturing a medicament for the
PT   treatment of HCV.
XX
PS   Example 2; Page 27; 79pp; English.
XX
CC   The present sequence is that of the NS5B protein of hepatitis C virus
CC   (HCV). The sequence is the translation sequence of a polynucleotide in
CC   which codon usage was altered to resemble that of highly expressed human
CC   genes. HCV vaccines of the invention comprise a polynucleotide that
CC   encodes the HCV proteins Core, NS3, NS4B and NS5B, and does not encode
CC   the NS4A and/or NS5A proteins. The proteins may be expressed as
CC   individual proteins or as fusion proteins. Preferred fusions include
CC   double fusions between NS4B and NS5B and between Core and NS3. The
CC   vaccines are useful for the treatment or prevention of an HCV infection.
XX
SQ   Sequence 592 AA;

Query Match   95.2%; Score 3014; DB 8; Length 592;
Best Local Similarity   96.1%; Pred. No. 1.8e-279;
Matches 569; Conservative 14; Mismatches   9; Indels   0; Gaps   0;

Qy   4 MSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQV 63
            ||||||||||||||||||||||||||| ||:||||||||||:|:||||||||||||||||
Db   1 MSMSYTWTGALITPCAAEESKLPINPLSNSLRHHNMVYATTSRSASLRQKKVTFDRLQV 60

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHI 123
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LDDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHI 120

Qy   124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 183
            ||||:|||||:||||||||||:|||||||||||||||||||||||||||||||||||||
Db   121 RSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 180

Qy   184 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
            ||||||||||||||||||||||||||||||||||| |||||||||||:|||||||||||
Db   181 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYGTRCFGSTVTESDIRVEESI 240

Qy   244 YQCCDLAPEARQAIRSLTERLYVGGMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 303
            ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   241 YQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 300

Qy   304 AAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYD 363
            |||:|||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   301 TAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYD 360

Qy   364 LELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTL 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 LELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTL 420

Qy   424 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 483
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 480

Qy   484 GEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPI 543
            ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   481 GEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPI 540

Qy   544 PAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGGIYLLPNR 595
            ||:||||||||||||||||||||||||||||| ||||||||||||||||||
Db   541 PAASQLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLLSVGGIYLLPNR 592

RESULT 15
AAY06423
ID   AAY06423 standard; protein; 3010 AA.
XX
AC   AAY06423;
XX
DT   20-MAR-2003 (revised)
DT   27-SEP-1999 (first entry)
XX
DE   Non-A, non-B hepatitis virus polypeptide.
XX
KW   Non-A, non-B hepatitis B virus; NANBV; antigen; infection; diagnosis;
KW   vaccine.
XX
OS   Non-A.
OS   non-B hepatitis virus.
XX
FH   Key   Location/Qualifiers
FT   Protein   1..115
FT   /note= "core protein"
FT   Protein   116..191
FT   /note= "matrix protein"
FT   Protein   192..389
FT   /note= "envelope protein"
FT   Protein   390..729
FT   /note= "NS1 protein"
FT   Protein   730..1006
FT   /note= "NS2 protein"
FT   Protein   1007..1615
FT   /note= "NS3 protein"
FT   Protein   1616..1862
FT   /note= "NS4a protein"
FT   Protein   1863..2013
FT   /note= "NS4b protein"
FT   Protein   2014..3010
FT   /note= "NS5 protein"
XX
PN   EP933426-A1.
XX
PD   04-AUG-1999.
XX
PF   28-DEC-1990; 99EP-00106005.
XX
PR   25-JUN-1990; 90JP-00167466.
PR   31-AUG-1990; 90JP-00230921.
PR   09-NOV-1990; 90JP-00305605.
PR   28-DEC-1990; 90EP-00314371.
XX
PA   (OSAU) UNIV OSAKA.
XX
PI   Okayama H, Fuke I, Mori C, Takamizawa A, Yoshida I;

Search completed: September 22, 2005, 14:47:11
Job time : 172 secs

```
XX
DR   WPI; 1999-407152/35.
DR   N-PSDB; AAX59394.
XX
PT   New hepatitis virus polypeptides, useful for diagnosing and treating
PT   hepatitis infections.
XX
PS   Claim 2; Fig 2(1)-(16); 56pp; English.
XX
CC   This sequence represents the non-A, non-B hepatitis virus (NANBV)
CC   polypeptide, as predicted from cDNA (see AAX59394) containing the entire
CC   open reading frame of the NANBV genome. To obtain this cDNA, NANBV RNAs
CC   were extracted directly from NANBV particles contained in whole blood of
CC   a patient having NANB hepatitis, or from a resected liver of a patient
CC   having NANB hepatitis and liver cancer. The RNA was then converted to
CC   double-stranded cDNA. A cDNA library was produced and screened using
CC   serum from a convalescent patient having acute NANB hepatitis and serum
CC   from a patient having chronic NANB hepatitis. The isolated cDNA allows
CC   recombinant production of NANBV antigen polypeptides in microbial or
CC   eukaryotic cell culture. The method provides the safe production of NANBV
CC   antigens with high purity on a large scale at low cost without the
CC   biohazard associated with multiplying virus in animals. Claimed NANBV
CC   nucleotide sequences are useful for the recombinant production of
CC   polypeptides useful as antigens for vaccines, and as diagnostic reagents.
CC   (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC   correct PR field.)
XX
SQ   Sequence 3010 AA;

Query Match          95.2%;   Score 3014;   DB 2;   Length 3010;
Best Local Similarity  96.6%;  Pred. No. 2.3e-278;
Matches  571;  Conservative  10;  Mismatches  10;  Indels  0;  Gaps  0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL  64
          |||||||||||||||||||||||||||||||||| ||| : ||||| ||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||| ||||||||| ||| ||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          |||||||||| ||| :|||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||||.|||||||||||| |||||||||| ::| |||||||||||||||||||||||||||
Db   2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||:||||||| |||:|||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||.||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||.|||||||||:|||||||||| :|||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRASLLSQGGRAATCGKYLFNWAVTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      September 22, 2005, 14:40:28 ; Search time 19 Seconds
                                (without alignments)
                                3043.488 Million cell updates/sec

Title:       US-10-712-479-2
Perfect score: 3167
Sequence:    1 MASMSMSYTWTGALITPCAA.........LLSVGVGIYLLPNRHHHHHH 601
Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:      283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_79:*
             1: pir1:*
             2: pir2:*
             3: pir3:*
             4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3016	95.2	3010	1	GNWVCJ	genome polyprotein
2	3016	95.2	3010	1	GNWVTC	genome polyprotein
3	3003	94.8	3010	1	A45573	genome polyprotein
4	2968	93.7	3010	1	GNWVTW	genome polyprotein
5	2944	93.0	3010	1	S18030	genome polyprotein
6	2798	88.3	3011	1	S40770	genome polyprotein
7	2764	87.3	3011	1	GNWVC3	genome polyprotein
8	2725	86.0	3011	1	GNWVCH	genome polyprotein
9	2511	79.3	3014	2	JC5620	genome polyprotein
10	2392	75.5	3033	2	GNWVJ8	genome polyprotein
11	2380	75.1	3033	2	JQ1303	genome polyprotein
12	1886	59.6	365	2	JQ0879	NS5 protein - hepa
13	1750	55.3	365	2	JQ0880	NS5 protein - hepa
14	1471.5	46.5	874	2	JQ0883	genome polyprotein
15	1442	45.5	874	2	JQ0881	genome polyprotein
16	951	30.0	189	2	S60587	non-structural pro
17	774	24.4	284	2	PS0104	genome polyprotein
18	744	23.5	259	2	PS0102	genome polyprotein
19	719	22.7	156	2	D39109	genome polyprotein
20	667	21.1	135	2	S44214	genome polyprotein
21	654	20.7	3005	2	T08841	polyprotein - dour
22	636	20.1	1435	2	T01075	polyprotein - hepa
23	629	19.9	2970	2	T08839	polyprotein - marm
24	569	18.0	113	2	PC1278	NS5 protein - hepa
25	563	17.8	113	2	PC1277	NS5 protein - hepa
26	562	17.7	113	2	PC1274	NS5 protein - hepa
27	562	17.7	113	2	PC1275	NS5 protein - hepa
28	560	17.7	113	2	PC1276	NS5 protein - hepa
29	513	16.2	113	2	PC1279	NS5 protein - hepa
30	426	13.5	113	2	PC1282	NS5 protein - hepa
31	426	13.5	113	2	PQ0275	polyprotein - hepa
32	425	13.4	113	2	PQ0276	polyprotein - hepa
33	424	13.4	113	2	PC1281	NS5 protein - hepa
34	422	13.3	113	2	PC1283	NS5 protein - hepa
35	409	12.9	113	2	PC1280	NS5 protein - hepa
36	408	12.9	113	2	PQ0273	polyprotein - hepa
37	406	12.8	113	2	PQ0274	polyprotein - hepa
38	402	12.7	78	2	PS0085	hypothetical prote
39	336	10.6	266	2	PQ0393	genome polyprotein
40	300	9.5	132	2	PQ0396	genome polyprotein
41	298	9.4	132	2	PQ0394	genome polyprotein
42	296	9.3	60	2	S44215	genome polyprotein
43	264	8.3	71	2	PC1300	genome polyprotein
44	261	8.2	71	2	PC1302	genome polyprotein
45	260	8.2	71	2	PC1301	genome polyprotein

ALIGNMENTS

```
RESULT 1
GNWVCJ
genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural
  protein NS4a; nonstructural protein NS4b; envelope protein NS5
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A39253; PS0086
R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimotc
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients v
A;Reference number: A39253; MUID:91088550; PMID:2175903
A;Accession: A39253
A;Molecule type: genomic RNA
A;Residues: 1-3010 <KAT>
A;Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
A;Reference number: PS0085
A;Accession: PS0086
A;Molecule type: genomic RNA
A;Residues: 2650-2707 <KA2>
A;Experimental source: Japanese isolate
C;Comment: The cleavage sites of this polyprotein have not been determined.
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2:

Query Match        95.2%;   Score 3016;  DB 1;  Length 3010;
Best Local Similarity  96.3%;   Pred. No. 3.8e-223;
Matches 569;  Conservative 12;  Mismatches 10;  Indels 0;  Gaps 0;

Qy     5 SMSYTWTGALITPCAAEBSKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           |||||||||||||||||||||||||||||||| ::|||||||||||||||||||||||
Db  2420 SMSYTWTGALITPCAAEBSKLPINPLSNSLLRHSMVYSTTSRSASLRQKKVTFDRLQVL 2479

Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||| :|||| ::|||||||||||||||||||||||||||||||
Db  2480 DDHYRDVLKEMKAKASTVKARLLSIEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 2539
```

```
Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSRAVNHIR 124
          |||||||||||||||:|||||||||||||||||||||:|||||||||||||:|||:|||
Db  2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSKAVNHIH 2539

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2540 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||||:||||||:||||||||||||||||||||||||||:||||||
Db  2600 TLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESLY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||::
Db  2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||:||||||||||||||||||||||:|||||||||
Db  2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 3
A45573

genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98)
    protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: A45573
R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, M.;
    Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier; se
A;Reference number: A45573; MUID:92295714; PMID:1318627
A;Accession: A45573
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3010 <TAN>
A;Cross-references: UNIPROT:Q00269; GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PIDN:
A;Experimental source: HCV-JT
A;Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match   94.8%; Score 3003; DB 1; Length 3010;

RESULT 2
GNWTC

genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
    protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A38465
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
    J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from human
A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Accession: A38465
A;Molecule type: genomic RNA
A;Residues: 1-3010 <TAK>
A;Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g3297
A;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match   95.2%; Score 3016; DB 1; Length 3010;
Best Local Similarity   96.6%;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

```
Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          ||||||||||||||||||||||||||||||||||||:||||||:|||||||||||||||
Pb  2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479
```

```
Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db  2540 SVWEDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRTEESIY 244
          ||||||||||||||||||||||||||||||||:||||||||||||||||||||||:||||
Db  2600 TLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRTEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||:|
Db  2660 QCCDLAPEARQAIKSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||:|||| |||:|||||||||||||||||||||
Db  2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db  2780 ELITSCSSNVSVAHDASGKRVYVLTRDPTTPLARAAWETVRHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||:||||||||||||||||||||||:|||||||||
Db  2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2960 AASQLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

Best Local Similarity 95.9%; Pred. No. 3.8e-222;
Matches 567; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

```
Qy       5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           |||||||||||||||||||||||||||||||||||||||:|||||||:|||||||||||||
Db    2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASKFGYGAKDVRSLSSRAVNHIR 2479

Qy      65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db    2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSGKAINHIR 2539

Qy     125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db    2540 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy     185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           |||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||
Db    2600 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 2659

Qy     245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||:|
Db    2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy     305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2720 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPLPQPEYDQ 2779

Qy     365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWATARHTPVNSWLGNIIMYAPTLW 2839

Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db    2900 EINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGRYLFNWAVKTKLKLTPIP 2959

Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||:||:|||||||||||||||||||||||||||||||||||||||||||
Db    2960 AASQLDLSKWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 4
GNWVTW
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu...
   protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A;Reference number: A40244; MUID:92230206; PMID:1314449
A;Accession: A40244
A;Molecule type: genomic RNA
A;Residues: 1-3010 <CHE>
A;Cross-references: UNIPROT:P29846; GB:M84754
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura...
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <BPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250;305,325,417,423,430,448,532,540,556,623,645,1213,1255,2041,207...

Query Match 93.7%; Score 2968; DB 1; Length 3010;
Best Local Similarity 94.9%; Pred. No. 1.9e-219;
Matches 561; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

```
Qy       5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           ||||||||||||||||||||||||||||||||||||||:|||||||:|||||||||||||
Db    2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTIDRLQVL 2479

Qy      65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||||||||||||||||||||||||||||||||:|:|||:|||||
Db    2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSGKAINHIR 2539

Qy     125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||:|||||:||||||||||||||||||||||||||||||||||||||||||||
Db    2540 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy     185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||||||||||||||||:|||||||||||||:|||||||||||||||||:|||||||||
Db    2600 TLPQAVMGSSYGFQYSPQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 2659

Qy     245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||:
Db    2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy     305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           |||||||||||||||:|||||||||||||||||||||||||||||||||||||:||||||
Db    2720 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPLPQPEYDQ 2779

Qy     365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           |||||||||||||||||||||||||||||:||||||:|||||||||||||||||||||
Db    2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWATARHTPVNSWLGNIIMYAPTLW 2839

Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||:|||||||||||||||||||||:|||||||||:|||||||||
Db    2900 EINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGRYLFNWAVKTKLKLTPIP 2959

Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||:||:|||||||||||||||||||||||||||||||||||||||||||
Db    2960 AASQLDLSKWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 5
S18030
genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructur...
   protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Variety: isolate JK1
C;Date: 19-May-2000 #sequence revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: S18030; S33570; A48332; S18029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A;Description: A whole genome of hepatitis C virus cDNA was isolated from a single patien...
A;Reference number: S18028
A;Accession: S18030
A;Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
A;Experimental source: isolate JK1 from an individual
R;Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A;Title: Sequence analysis of putative structural regions of hepatitis C virus isolated f...
A;Reference number: A48332; MUID:93119270; PMID:8380322
A;Accession: S33570
A;Molecule type: genomic RNA

A;Residues: 1-547,'T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
A;Cross-references: EMBL:X61591
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320
     as Trp, and TTC for residue 771 as Ser
A;Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein E #status predicted <EPM>
F;192-389/Product: major envelope protein NS1 #status predicted <MEE>
F;390-729/Product: nonstructural protein NS2 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As

Query Match   93.0%;   Score 2944;   DB 1;   Length 3010;
Best Local Similarity   93.9%;   Pred. No. 1.3e-217;
Matches   559;   Conservative   17;   Mismatches   19;   Indels   0;   Gaps   0;

Qy   1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
Db 2416 VACCSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSAGLRQKKVTFDR 2475

Qy   61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
Db 2476 LQVPDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKAV 2535

Qy   121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVPDLGVRVCEKMALY 180
Db 2536 NHIHSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVPELGVRVCEKMALY 2595

Qy   181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE 240
Db 2596 DVVSTLPQAVMGSSYGFQYSPGQRVEFLVQAWKSKRTPMGFSYDTRCFDSTVTESDIRVE 2655

Qy   241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300
Db 2656 ESIYQCCDLDPQARVAIRSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCY 2715

Qy   301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
Db 2716 LKASAACRAAKLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQP 2775

Qy   361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420
Db 2776 EYDLELITSCSSNVSVAHDGTGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYA 2835

Qy   421 PTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
Db 2836 PTLWARMLMTHFFSILLAQEQLEKALDCQIYGATYFIEPLDLPQIIQRLHGLSAFSLHS 2895

Qy   481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540
Db 2896 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKL 2955

Qy   541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db 2956 TPIPAASQLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 3010

RESULT 6
S40770
genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
   protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus

C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: S40770; PC1285
R;Okamoto, H.
   submitted to the EMBL Data Library, March 1992
A;Reference number: S40770
A;Accession: S40770
A;Molecule type: genomic RNA
A;Residues: 1-3011 <OKA>
A;Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g2215i
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Accession: PC1285
A;Molecule type: genomic RNA
A;Residues: 1-513 <OK2>
A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A;Experimental source: isolate HC-J1
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match   88.3%;   Score 2798;   DB 1;   Length 3011;
Best Local Similarity   88.8%;   Pred. No. 2.3e-206;
Matches   525;   Conservative   28;   Mismatches   38;   Indels   0;   Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
Db 2421 SMSYTWTGALITPCAAEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 2480

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
Db 2481 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVNHIN 2540

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db 2541 SVWKDLLEDSVTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2600

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
Db 2601 KLPPAVMGSSYGFQYSPGQRVEFLVQAWKSKRTPMGFSYDTRCFDSTVTESDIRTEEAIY 2660

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
Db 2661 QCCDLDPQARVAIRSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKAR 2720

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
Db 2721 AACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDL 2780

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
Db 2781 ELITSCSSNVSVAHDGTGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLW 2840

Qy   425 ARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Db 2841 ARMLMTHFFSVLIARDLEQALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPG 2900

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
Db 2901 EINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFNWAVRTKLLTFIA 2960

```
Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVIYLLPNR   595
        ||||||||||||:||||||||||||||:||:|||||||:||||||||||
Db 2961 AAGRLDLSGWFTAGYSGGDIYHSVSHARPRWFWFCLLLLAAGVGIYLLPNR 3011
```

RESULT 7
GNWVC3
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A39166; PQ0403; PQ0404
R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi
 Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
A;Accession: A39166
A;Molecule type: mRNA
A;Residues: 1-3011 <CHO>
A;Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L
 J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A;Reference number: PQ0403; MUID:92268871; PMID:1316939
A;Accession: PQ0403
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CHA>
A;Cross-references: DDBJ:D10128
A;Experimental source: isolates E-b16
A;Accession: PQ0404
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CH2>
A;Experimental source: isolates E-b17
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <MEE>
F;192-389/Product: major envelope protein E #status predicted <EPM>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,22

```
Query Match    87.3%;  Score 2764;  DB 1;  Length 3011;
Best Local Similarity 87.5%;  Pred. No. 9.6e-204;
Matches  517;  Conservative  32;  Mismatches  42;  Indels  0;  Gaps  0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
        |||||:||||:|||||||:||||||||||||||||||||||||||:  |:||||||||||
Db 2421 SMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRKKKVTFDRLQVL 2480

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
        |:||:|||||:||:|||:|||:||||||| :|.|||||||||||||||||: ::|.|.|
Db 2481 DSHYQDVLKEVKAAASKVKANLLSVEEACSLAPPHSAKSKFGYGAKDVRCHARKAVAHIN 2540

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
        ||||||||| :|||:|||||||||||||||||||||||||||||||||||||||||||||
Db 2541 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 2600

Qy  185 TLPQAVMGSSYGFQYSPGQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
        .||.||||||||||||||||||||| .||:||.|||||||||||||||||.|||.||:||
Db 2601 KLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIY 2660

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
        ||||| :|||.||.||||||||||||||.:||.|||||||||||||||||||||||:||
Db 2661 QCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKAR 2720
```

```
Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
        ||||||.|||||||||.||||||||||||.|||||||||.||||||||||||||||||||
Db 2721 AACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDL 2780

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
        |||||||||||||||.:|||||||||||||.||||||||||||||||||||||||:||||
Db 2781 ELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLW 2840

Qy  425 ARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG   484
        |||||||||:|:|.::||:.||||:|||||||||||||||:||:||||||||||||||||
Db 2841 ARMLMTHFFSVLIARDQLEQALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPG  2900

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
        |||||.|||||||||||||.|||||||||:||...||||||||||||||||||||||||.|
Db 2901 EINRVAACLRKLGVPPLRAWRHRARSVRARLLARGGRAAICGKYLFNWAVRTKLKLTPIA 2960

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVIYLLPNR            595
        ||.:||||||||:||||||||||||:||:||||||:||:  ||||||||||||||||||||
Db 2961 AAGQLDLSGWFTAGYSGGDIYHSVSHARPRWIWFCLLLLAAGVGIYLLPNR         3011
```

RESULT 8
GNWVCH
genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36814; A41546
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
 submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: cc
A;Reference number: A36814
A;Accession: A36814
A;Molecule type: genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: UNIPROT:P27958; GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
 Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: comparis
A;Reference number: A41546; MUID:92052256; PMID:1658800
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <MEE>
F;192-389/Product: major envelope protein E #status predicted <EPM>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

```
Query Match    86.0%;  Score 2725;  DB 1;  Length 3011;
Best Local Similarity 86.6%;  Pred. No. 9.6e-201;
Matches  512;  Conservative  34;  Mismatches  45;  Indels  0;  Gaps  0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
        |||||:||||:|||||||:||||||||||||||||||||||||||:  |:||||||||||
Db 2421 SMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRKKKVTFDRLQVL 2480

Qy   65 DDHYRDVLKEMKAKASTVKAKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR          124
        |:||:|||||:||:|||:|||  :|||||||||||||||||: ::|.|.|
Db 2481 DSHYQDVLKEVKAAASKVKANLLSVEEACSLAPPHSAKSKFGYGAKDVRCHARKAVAHIN 2540
```

```
Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
        |||||||||||||| |||| |||| ||| |||||||||||||||| | ||||||||||||
Db 2424 SMSYTWTGALITPCSAEEKLPINPLSNTLLRHHNLVYSTTSRSSAGLRQKKVTFDRLQVL 2483

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
        |||||:|::.:||.:::.|||:||..:||.||||||:||:|||||:|||||..:|.:.:
Db 2484 DDHYREVVDEMKDRLASKVKARLLPEEACGLTPPHSARSKYGYGAKEVRSLDKKALKHIE 2543

Qy  125 SVWKDLLEDTDPIQTTIMAKQNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
        .:|.||| |:. | ||||||||||| |:|.||.|||||||| |||||||||| |||||.
Db 2544 GVWQDLLDDSDTPLPTTIMAKQNEVFAVEPSKGGKKPARLIVVPDLGVRVCEKRALYDVAQ 2603

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
        .||:|:||.||||||||.||::||:..||:||.||.|||||||||||...:||: |EESIY
Db 2604 KLPTALMGPSYGFQYSPAQRVDFLLKAWKSKKIPMAFSYDTRCFDSTTEHDIMTEESIY 2663

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
        |.|||.|||:||||||:.|||.|||||||||:.|||||||||||:|||. ||:||||:||
Db 2664 QSCDLQPEARVAIRSLTQRLYGGPMTNSKGQQCGYRRCRASGVFTTSMGNTMTCYIKAL 2723

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
        |.|||||::||.:||.|||||.|||| ||:||:||||.|||||||||||||||||.|.||||
Db 2724 ASCRAAKLRDCTLLVCGDDLVAICESQGTHEDEASLRAFTEAMTRYSAPPGDPPVPAYDL 2783

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
        ||:|||||||||| ||||:|:|||||||| |||::|||||:|:.|||||||||||||||||
Db 2784 ELVTSCSSNVSVARDASGNRIYYLTRDPQVPLAKAAWETAKHSPVNSWLGNIIMYAPTLW 2843

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
        ||::|||||||:| .:|||||:|.:|:||:. |.:.||||||:||:||||||||||||||:
Db 2844 ARIVLMTHFFSVLQSQEQLEKTLAFEMYGSVSVTPLDLPAIIQRLHGLSAFSLHSYSPS 2903

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLAQGGRAAICGILYLFNWAVKTKRKLTPLA  544
        |||||||||||||||||||:|||||:|||||||||||||||||.||||||||||||||||||
Db 2904 EINRVASCLRKLGVPPLRAWRHRARAVRAKLIAQGGRAAICGILYLFNWAVKTKRKLTPLA 2963

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
        .|.|||||.||.||.||||||||:|||||||.:.||||||||||||.||.|
Db 2964 DADRLDLSSWFTVGAGGDIYHSMSRARPRNLLCLLLLSVGVGIFLLPAR 3014
```

RESULT 10
GNWVJ8
genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40250; PQ0397; PQ0559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to repo
A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
A;Molecule type: genomic RNA
A;Residues: 1-3033 <OKA>
A;Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; P
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L.
J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e>
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0397
A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DDBJ:D10134
A;Experimental source: isolate E-b12
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno,
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
A;Accession: PQ0559
A;Molecule type: mRNA
A;Residues: 2678-2729 <KAT>

```
Qy  125 SVWKDLLEDTDPIQTTIMAKQNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
        ||||||||| . | ||||||||||||||||||||||||||| |||||||||||||||||||
Db 2541 SVWKDLLEDSVTPIDTTIMAKQNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2600

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
        ||||||||||||||||| |||||||..||:||.|||.||||||||||||.:|||.EE:IY
Db 2601 TLPQAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGLSYDTRCFDSTVTESDIRTEEAIY 2660

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
        |||||.|:||.||:|||||||||||:||:.|:||||||||||.|||||||||||.|:||.
Db 2661 QCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASRVLTTSCGNTLTRYIKAR 2720

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
        ||||||.||||.:||..||||||||||.:||:||||||.||||||||||||||||||||||
Db 2721 AACRAAGLQDCTMLVVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDL 2780

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
        ||||||||||||||| :|:||||||||||.|||||||||||||||||||||||||||||||
Db 2781 ELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2840

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
        ||||||||||:|:|.:.||||:|.|:||||||||||||||:||:||||||||||||||||
Db 2841 ARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPG 2900

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
        ||||||.||||||||||| ||||| ||||:||::||:|||||||||||||||||||||||
Db 2901 EINRVAACLRKLGVPPLRAWRHRAWSVRARLLARGGKAAICGKYLFNWAVRTKLKLTPIT 2960

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
        ||.||||||||.||||||||||||:..|||||||:.||||:.|||||||||
Db 2961 AAGRLDLSGWFTAGYSGGDIYHSVSHARPRWFWFCLLLLAAGVGIYLLPNR 3011
```

RESULT 9
JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9233423
A;Experimental source: genotype 5a, which predominates in South Africa
A;Cross-references: UNIPROT:O39928; GB:Y13184
A;Note: the translation of the nucleotide sequence is not complete in this paper
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;388-408/Region: hypervariable #status predicted
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F;1008-1616/Product: hepacivirin #status predicted <NS3>
F;1231-1238/Region: nucleotide-binding motif A (P-loop)
F;1313-1318/Region: nucleotide-binding motif B
F;1317-1320/Region: DEXH motif
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F;2210-2249/Region: interferon sensitivity determining #status predicted

Query Match　79.3%; Score 2511; DB 1; Length 3014;
Best Local Similarity 78.8%; Pred. No. 2.9e-184;
Matches 466; Conservative 53; Mismatches 72; Indels 0; Gaps 0;

A;Cross-references: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <MEE>
F;192-389/Product: major envelope protein E #status predicted <NS1>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: hepacivirin #status predicted <NS3>
F;1011-1619/Product: hepacivirin #status predicted <NS3>
F;1234-1241/Region: nucleotide-binding motif A (P-loop)
F;1316-1321/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,23

Query Match   75.5%;   Score 2392;   DB 1;   Length 3033;
Best Local Similarity 75.5%;   Pred. No. 4.2e-175;
Matches 446; Conservative   59; Mismatches   86; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTPDRLQVL 64
            ||||||||||||||||:::||||||||:||| |||:|||||||||||:|||||||||||
Db   2443 SMSYSWTGALITPCGPEEEKLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTPDRVQVL 2502

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |:||::||:::||:||:|:::::|:||||||||||||||||:||||:||||||||||||
Db   2503 DAHYDSVLQDVKRAASKVSARLLTVERACALTPPHSAKSRYGFGAKEVRSLSRAVNHIR 2562

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVPDLGVRVCEKMALYDVVS 184
            ||:||||||:|:|:|||||||||||: |||:||||||||||||||||||||||||
Db   2563 SVWEDLLEDQHTPIDTTIMAKNEVFCIDPTKGGKKPARLIVPDLGVRVCEKMALYDIAQ 2622

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||:||||||||| :||||:||||||||||:| ||||||||||||||||:|||||
Db   2623 KLPKAIMGPSYGFQYSPAERVDFLLKAWGSKKDPMGFSYDTRCFDSTVTERDIRTEESIY 2682

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||:|::||||:||||||||||||||||||||:||||:||||||||||||||| ||:||
Db   2683 QACSLPQEARTVIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKAL 2742

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||| |||:::|::||||||||||||||||:|||||||||||||||||||||||:||||||
Db   2743 AACKAAGIIAPVMLVCGDDLVVISESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDL 2802

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            ||||||||||| ||||:|||||||||||||| |||||||||| || |||||||:|||||
Db   2803 ELITSCSSNVSVALDSRGRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIW 2862

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            ||:::||||||||| ||||:|||||:||||||||::||| ||||||||||||||||||
Db   2863 VRMVIMTHFFSILLAQDTLNQNLNFEMYGAVSVNPLDLPAIIERLHGLEAFSLHTYSPH 2922

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||| :|||||||| |||||:||||||:|||||||| ||| |||||||| |||||||||
Db   2923 ELSRVAATLRKLGAPPLRAWKSRARAVRASLIAQGARAAICGRYLFNWAVKTKLKLTPLP 2982

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            ||||||||||| ||||||||||| ||||| |||| |:||| |||||:||||
Db   2983 EASRLDLSGWFTVGAGGGDIYHSVSHARPRLLLLGLLLLFVGVGLFLLPAR 3033

RESULT 11
JQ1303
genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains:capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JQ1303
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hum

---

A;Reference number: JQ1303; MUID:92044440; PMID:1658196
A;Accession: JQ1303
A;Molecule type: genomic RNA
A;Residues: 1-3033 <OKA>
A;Cross-references: UNIPROT:P26660; GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651
A;Experimental source: isolate HC-J6 from a Japanese individual
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transme
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F;1011-1619/Product: hepacivirin #status predicted <NS3>
F;1316-1321/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,281

Query Match   75.1%;   Score 2380;   DB 1;   Length 3033;
Best Local Similarity 75.6%;   Pred. No. 3.5e-174;
Matches 447; Conservative   57; Mismatches   87; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTPDRLQVL 64
            |||||||||||||| :::||||||:|||| |||:|||||||||||:|||||||||||
Db   2443 SMSYSWTGALITPCSPEEEKLPINPLSNSLLRYHNKVYCTTSKSASLRAKKVTPDRMQAL 2502

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIK 124
            |:||:|||::||:||:|:::::|:|||||||||||||||:||||||:||||||||||||
Db   2503 DAHYDSVLKDIKLAASKVTARLLTLERACQLTPPHSARSKYGFGAKEVRSLSGRAVNHIK 2562

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVPDLGVRVCEKMALYDVVS 184
            ||:|||||||:|:|:|||||||||| :|| ||:||||||||||||||||||||||
Db   2563 SVWKDLLEDTQTPIPTTIMAKNEVFCVDPTKGGKKAARLIVPDLGVRVCEKMALYDITQ 2622

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||:||||||||| :||||:||||||||||:||||||||||||||||||:|||||
Db   2623 KLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKDPMGFSYDTRCFDSTVTERDIRTEESIY 2682

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            ||||||||||:||||||||||||||||||||:||||:||||||||||||:||||||||
Db   2683 RACSLPEEAHTAIHSLTERLYVGGPMFNSKGQTCGYRRCRASGVLTTSMGNTTCYVKAL 2742

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            |||:|||||:|:||||||||||||||:||:|||:||||||||||||||||||:||||||
Db   2743 AACKAAGIIAPTMLVCGDDLVVISESQSTEEDERNLRAFTEAMTRYSAPPGDPRPEYDL 2802

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||||||| ||||||||| ||||||||||:|||||
Db   2803 ELITSCSSNVSVALGPQGRRRYLTRDPTTPIARAAWETVRHSPVNSWLGNIIQYAPTIW 2862

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            ||:::||||||||| |||||:|||||:||||||||::||| ||||||||||||||||||
Db   2863 ARMVLMTHFFSILMAQDTLDQNLNFEMYGAVSVSPLDLPAIIERLHGLDAFSLHTYTPH 2922

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||| :|||||||| |||||:||||||||:||||||| ||| |||||||| |||||||||
Db   2923 ELTRVASALRKLGAPPLRAWKSRARAVRASLISRGRAAVCGRYLFNWAVKTKLKLTPLP 2982

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            |||||||||||| ||||||||||| |||||| |:||:||| ||||||||||
Db   2983 EARLLDLSSWFTVGAGGGDIYHSVSRARPRLLLLGLLLLFVGVGLFLLPAR 3033

RESULT 12
JQ0879
NS5 protein - hepatitis C virus (strain J4) (fragment)
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0879
R;Okamoto, H.

```
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0879
A;Accession: JQ0879
A;Molecule type: genomic RNA
A;Residues: 1-365 <OKA>
A;Cross-references: UNIPROT:Q81717
A;Experimental source: strain J4
C;Superfamily: hepatitis C virus genome polyprotein

  Query Match    59.6%;   Score 1886;   DB 2;   Length 365;
  Best Local Similarity  96.7%;   Pred. No. 2.1e-137;
  Matches  353;   Conservative  9;   Mismatches  3;   Indels  0;   Gaps  0;

Qy  231  TVTENDIRVEBSIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLT  290
              |||||||||||||||||||||||:|:|||||||||:|:||||||||||||||||||||
Db    1  TVTESDIRVEBSIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLT  60

Qy  291  TSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRY  350
         ||||||||||||:||:||||||||||||||||||||||||||||||||||||||||||||
Db   61  TSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRVFTEAMTRY  120

Qy  351  SAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVN  410
         ||||||||||||||||||||||||||||||||||||||||:|:||||||||||||||||||
Db  121  SAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTIPLARAAWETARHTPVN  180

Qy  411  SWLGNIIMYAPTLWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIERL  470
         ||||||||||||||||||||||||||||||||:||::|||||||||:||||||||||:||
Db  181  SWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACYSIEPLDLPPIIQRL  240

Qy  471  HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLF  530
         ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db  241  HGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLF  300

Qy  531  NWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIY  590
         ||||||||||||||||:|||||||||||||||||||||||||||||||:|||||:|||||
Db  301  NWAVRTKLKLTPIAAAGRLDLSGWFTAGYSGGDIYHSVSHARPRWFWFCLLLLAAGVGIY  360

Qy  591  LLPNR  595
         |||||
Db  361  LLPNR  365
```

```
RESULT 13
JQ0880
NS5 protein - hepatitis C virus (strain J1) (fragment)
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0880
R;Okamoto, H.
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0880
A;Accession: JQ0880
A;Molecule type: genomic RNA
A;Residues: 1-365 <OKA>
A;Cross-references: UNIPROT:Q81716
A;Experimental source: strain J1
C;Superfamily: hepatitis C virus genome polyprotein

  Query Match    55.3%;   Score 1750;   DB 2;   Length 365;
  Best Local Similarity  88.2%;   Pred. No. 6.1e-127;
  Matches  322;   Conservative  23;   Mismatches  20;   Indels  0;   Gaps  0;

Qy  231  TVTENDIRVEBSIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLT  290
              |||||||||||||||||||||||:||||||||:|:|||:|||||||||||||||||||
Db    1  TVTESDIRTEBAIYQCCDLDQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLT  60

Qy  291  TSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRY  350
         ||||||||||||:||:||||:|||:||||||||||||||||:||:|||||||||||||||
Db   61  TSCGNTLTCYIKAQAACRAAGLRDCTMLVCGDDLVVICESGVQEDAASLRAFTEAMTRY  120

Qy  351  SAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVN  410
```

```
RESULT 14
JQ0883
genome polyprotein - hepatitis C virus (strain J7) (fragments)
N;Contains: NS5 protein
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0883
R;Okamoto, H.
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0879
A;Accession: JQ0883
A;Molecule type: genomic RNA
A;Residues: 1-874 <OKA>
A;Cross-references: UNIPROT:Q7LZY6
A;Experimental source: strain J7
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
F;510-874/Product: NS5 protein (fragment) #status predicted <NS5>

  Query Match    46.5%;   Score 1471.5;   DB 2;   Length 874;
  Best Local Similarity  70.0%;   Pred. No. 5.6e-105;
  Matches  282;   Conservative  40;   Mismatches  64;   Indels  17;   Gaps  3;

Qy  195  YGFQYSPKQRVEFLVNTWKAKKCPMGF-SYDTRCFD-STVTENDIRVEBSIYQCCDLAPE  252
         |||||||||||||              :|||||:||||||||||||||||||||||||
Db  487  YCWHYPPK------------------PCGIVSAKTVCGPVYTVTERDIRTEESIYQACSLPQE  531

Qy  253  ARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKL  312
         ||:||||||||||||||||||||||||||||||||||:||||||||||:||||:||||:||
Db  532  ARTAIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVFTSIGNTMTCYIKALLAACKAAGI  591

Qy  313  QDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSS  372
         :||||||:|||||||||:|||:||:||||||:|||||||:|||||||||||||||||||||
Db  592  KDPIMLVCGDDLVVISSSQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSCSS  651

Qy  373  NVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMTH  432
         ||||:|||:|||||||||||||:|:|||||||||||||||||||||:||:||||||:||:|
Db  652  NVSVALDPRGRRRYYLTRDPTTPISRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMTH  711

Qy  433  FFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASC  492
         ||:||||:|:||:||||||||:||||:||||||||:|||||||:||||||||||||||:||
Db  712  FFNILLAQDTLNQNLNFEMYGAVYSVNPLDLPAIIERLHGLDAFSLHTYSPHELSRVTAAT  771

Qy  493  LRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAASRLDLS  552
         |||||:||||:|||:||:||:||||||:||||||||||||||||||||||||:|:|||:|||
Db  772  LRKLGAPPLRAWKSRARAVRASLIAQGGRAAICGRYLFNWAVKTKLKLTPLPEAARLDLS  831

Qy  553  GWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
         |||:|||:|||||||:|||||||:|||||||:||||||||||||
Db  832  GWFTVGAGGDIFHSVSHARPRLLLCLLLLSVGVGIFLLPAR  874
```

RESULT 15
JQ0881
genome polyprotein - hepatitis C virus (strain J6) (fragments)
N;Contains: NS5 protein
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0881
R;Okamoto, H.
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0879
A;Accession: JQ0881
A;Molecule type: genomic RNA
A;Residues: 1-874 <OKA>
A;Cross-references: UNIPROT:Q7LZY5
A;Experimental source: strain J6
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
F;510-874/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match   45.5%;   Score 1442; DB 2;   Length 874;
Best Local Similarity   74.2%;   Pred. No. 1e-102;
Matches 271;   Conservative 58;   Mismatches 36;   Indels   0;   Gaps   0;

Qy   231 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLT 290
             |||:|||| ||||||: :| : ||::| |||||||||||:||||| ||||||||||||||
Db   510 TVTERDIRTEESIYRACSLPEEAHTAIHSLTERLVVGGPMFNSKGQTCGYRRCRASGVLT 569

Qy   291 TSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRY 350
             |||||:|| ||||:| ||:: |||||| |||||||||: :| |||||||||||||||||
Db   570 TSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQGTEEDERNLRAFTEAMTRY 629

Qy   351 SAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVN 410
             ||||||||: |||||||||||||||||:||| :|: |||||||||:|||||:||||||
Db   630 SAPPGDPPRPEYDLELITSCSSNVSVALGPQGRRRYLTRDPTTPIARAAWETVRHSPVN 689

Qy   411 SWLGNIIMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERL 470
             |||||||: ||||||||||||||||::| :|::|: ||::|:|| ||:||||||| |||
Db   690 SWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLNFEMYGAVVSVSPLDLPAIIERL 749

Qy   471 HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLF 530
             |||:||||:| |: :|| :||: | ||: |: |||:|||||||||||||||:||||||||
Db   750 HGLDAFSLHTYTPHELTRVASALRKLGAPPLRAWKSRARAVRASLISRGGRAAVCGRYLF 809

Qy   531 NWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVVGIY 590
             |||:|||||||| ||::|| |||| ||||| ||||||||||||:|||:||||||||::
Db   810 NWAVKTKLKLTPLPEARLLDLSSWFTVGAGGGDIYHSVSRARPRLLLLGLLLLFVGVGLF 869

Qy   591 LLPNR 595
             ||||
Db   870 LLPAR 874

Search completed: September 22, 2005, 14:48:09
Job time : 22 secs

This Page Blank (uspto)

```
              GenCore version 5.1.6
      Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    September 22, 2005, 14:36:26 ; Search time 78 Seconds
                                      (without alignments)
                          3945.636 Million cell updates/sec

Title:     US-10-712-479-2
Perfect score: 3167
Sequence:  1 MASMSMSYTWTGALITPCAA.........LLSVGVGIYLLPNRHHHHH 601

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:  1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:       1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : UniProt 03:*
           1: uniProt_sprot:*
           2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3040	96.0	3010	2	Q9DTE6	Q9dte6 hepatitis c
2	3040	96.0	3010	2	Q9J3IO	Q9j3iO hepatitis c
3	3039	96.0	3010	2	Q93077	Q93077 hepatitis c
4	3038	95.9	3010	2	Q9DTE9	Q9dte9 hepatitis c
5	3034	95.8	3010	2	Q81760	Q81760 hepatitis c
6	3030	95.7	3010	2	Q02829	Q02829 hepatitis c
7	3027	95.6	3010	2	Q68285	Q68285 hepatitis c
8	3027	95.6	3010	2	Q81825	Q81825 hepatitis c
9	3026	95.5	3010	2	Q9J3H0	Q9j3hO hepatitis c
10	3025	95.5	3010	2	Q9QIX6	Q9qix6 hepatitis c
11	3024	95.5	3010	2	Q9QIX5	Q9qix5 hepatitis c
12	3022	95.4	3010	2	P89966	P89966 hepatitis c
13	3022	95.4	3014	2	Q9DTE0	Q9dte0 hepatitis c
14	3021	95.4	3010	2	O09796	O09796 hepatitis c
15	3020	95.4	591	2	O39930	O39930 hepatitis c
16	3020	95.4	591	2	Q80MG2	Q80mg2 hepatitis c
17	3019	95.3	591	2	Q80MJ2	Q80mj2 hepatitis c
18	3018	95.3	591	2	Q9JGN9	Q9jgn9 hepatitis c
19	3018	95.3	591	2	Q68826	Q68826 hepatitis c
20	3018	95.3	3010	2	Q807P3	Q807p3 hepatitis c
21	3017	95.3	591	2	Q80MJ0	Q80mj0 hepatitis c
22	3016	95.2	2284	2	Q81817	Q81817 hepatitis c
23	3016	95.2	3010	1	POLG_HCVBK	P26663 h genome po
24	3016	95.2	3010	1	POLG_HCVJA	P26662 h genome po
25	3016	95.2	3010	2	Q9DTE4	Q9dte4 hepatitis c
26	3015	95.2	591	2	Q80MH3	Q80mh3 hepatitis c
27	3015	95.2	591	2	Q80MI0	Q80mi0 hepatitis c
28	3015	95.2	3010	2	092969	092969 hepatitis c
29	3015	95.2	3010	2	092971	092971 hepatitis c
30	3015	95.2	3010	2	092972	092972 hepatitis c
31	3015	95.2	3010	2	Q9J3H9	Q9j3h9 hepatitis c
32	3012	95.1	3010	2	Q99AU2	Q99au2 hepatitis c
33	3011	95.1	3009	2	Q9J3F9	Q9j3f9 hepatitis c
34	3009	95.0	3005	2	Q02828	Q02828 hepatitis c
35	3008	95.0	3009	2	Q9DTD9	Q9dtd9 hepatitis c
36	3007	94.9	3008	2	P90194	P90194 hepatitis c
37	3007	94.9	3007	2	Q9J3H7	Q9j3h7 hepatitis c
38	3006	94.9	3007	2	P90193	P90193 hepatitis c
39	3005	94.9	3005	2	Q80MH6	Q80mh6 hepatitis c
40	3005	94.9	3010	2	Q9WMX2	Q9wmx2 hepatitis c
41	3004	94.9	3004	2	Q80MH5	Q80mh5 hepatitis c
42	3004	94.9	3004	2	092970	092970 hepatitis c
43	3003	94.8	3010	1	POLG_HCVJT	POLG HCVJT h genome po
44	3002	94.8	3003	2	Q9QIY3	Q9qiy3 hepatitis c
45	3002	94.8	3002	2	Q9QIY4	Q9qiy4 hepatitis c

```
                          ALIGNMENTS

RESULT 1
Q9DTE6
ID   Q9DTE6      PRELIMINARY;      PRT;   3010 AA.
AC   Q9DTE6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Serum;
RA   Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA   Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA   Mishiro S.;
RT   "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT   with hepatocellular carcinoma: the 'progression score' revisited.";
RL   Hepatol. Res. 20:161-171(2001).
DR   EMBL; AB049091; BAB18804.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q81755; 1DXP.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
```

```
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD12;
RA   Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF207753; AAF65943.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE 3010 AA; 326693 MW; 074098DB305AF1A9 CRC64;

Query Match          96.0%;  Score 3040;  DB 2;  Length 3010;
Best Local Similarity   97.3%;  Pred. No. 6.3e-228;
Matches  575;  Conservative   9;  Mismatches   7;  Indels   0;  Gaps   0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||||||||||| || ||||||| ||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSACQRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||| ||||||||||||| |||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 2539
```

```
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
KW   Hydrolase; Nonstructural protein; Polyprotein; Transmembrane.
SQ   SEQUENCE 3010 AA; 327043 MW; 3807DC687968A4C95 CRC64;

Query Match          96.0%;  Score 3040;  DB 2;  Length 3010;
Best Local Similarity   97.1%;  Pred. No. 6.3e-228;
Matches  574;  Conservative  11;  Mismatches   6;  Indels   0;  Gaps   0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||||||||||| || ||||||| ||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||| ||||||||||||| |||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          |||||||||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||  ||||||| ||  |||||||||||||||||||||||||||||
Db   2600 TLPQAVMGSSYGFQYSPGQRVEFLVKTWKSKKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||| |
Db   2660 QCCDLAPEARQAIRSLTERLYIGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||| |||| ||| ||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLIRDPTTPLSRAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASQDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

```
RESULT 2
Q9J3I0
ID   Q9J3I0       PRELIMINARY;  PRT;  3010 AA.
AC   Q9J3I0;
DT   01-OCT-2000  (TrEMBLrel. 15, Created)
DT   01-OCT-2000  (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004  (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
```

```
Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
               ||||||||||  :||| |||| ||||||||||||||||||||||||||||||||||||||
Db   2540  SVWKDLLEDTETPINTTVMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
               ||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||
Db   2600  TLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
               |||||||||||||||||||||||||||||  :||| ||||||||||||||||||||||
Db   2660  QCCDLAPEARQVIRSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESSAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
               |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   2720  AACRAAKLQDCTMLVCGDDLVVICDSSAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2900  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy    545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

```
RESULT 3
093077
ID   093077     PRELIMINARY;      PRT;   3010 AA.
AC   093077;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9812498; PubMed=9462666; DOI=10.1002/hep.510270242;
RA   Aizaki H., Aoki Y., Harada T., Ishii K., Suzuki T., Nagamori S.,
RA   Toda G., Matsuura Y., Miyamura T.;
RT   "Full-length complementary DNA of hepatitis C virus genome from an
RT   infectious blood sample.";
RL   Hepatology 27:621-627(1998).
DR   EMBL; D89815; BAA25076.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; CYTOCHROME; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;   327352 MW;   888BA102A733390 CRC64;
```

```
Query Match             96.0%;
Best Local Similarity   97.6%;      Score 3039;  DB 2;   Length 3010;
Matches  577;  Conservative   7;  Pred. No. 7.5e-228;
                                    Mismatches   7;   Indels    0;   Gaps    0;

Qy      5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
               |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   2420  SMSYTWTGALITPCAAEESKLPINALSNPLLRHHNMVYSTTSRSASLRQKKVTFDRMQVL 2479

Qy     65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2480  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 2539

Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 2719

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESSAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720  AACRAAKLQDCTMLVNGDDLVVICESSAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
               ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2780  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840  ARMILMTHFFSILLAQEQLEKALDCQIYGATYSIEPLDLPQIQRLHGLSAFSLHSYSPG 2899

Qy    485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db   2900  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959
```

```
Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 595
         |:||||||||||||||||||||||||||||||| ||||||||||||||||
Db  2960 EASQLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 3010
```

RESULT 4
Q9DTE9
```
ID   Q9DTE9           PRELIMINARY;      PRT;   3010 AA.
AC   Q9DTE9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Serum;
RA   Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA   Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA   Mishiro S.;
RT   "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT   with hepatocellular carcinoma: the 'progression score' revisited.";
RL   Hepatol. Res. 20:161-171(2001).
DR   EMBL; AB049088; BAB18801.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
     HSSP; Q8IYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
QR   Pfam; PF00998; Viral_RdRp; 1.
```

```
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE  3010 AA;  326780 MW;  668CFFEA5FEC3658 CRC64;
```

```
Query Match             95.9%;  Score 3038;  DB 2;  Length 3010;
Best Local Similarity   96.8%;  Pred. No. 9e-228;
Matches  572;  Conservative  13;  Mismatches  6;  Indels  0;  Gaps  0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
         |||||||||||||||||||||||||||||||||||::|| :|||||:||||||||||||||
Db  2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMIYATTSRSAGLRQKKVTFDRLQVL 2479

Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
         ||||||||||||||||||||||||||||||||||||||||||||||||:|||:|:|||||
Db  2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 2539

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
         |||||||||:::||:||||||||:|||||||||||||||||||||||||||||||||||||
Db  2540 SVWKDLLEDSETPIDTTIMAKNEIFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
         |||||||||||||||||||:|||||||||:||:||||||||||||||||||||||:|||||
Db  2600 TLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRIEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
         |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||:
Db  2660 QCCDLAPEARQAIRSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
         |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db  2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRAFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
         |||||||||||||||||||||||||||||||||||||||:|||||||:||||||||||||
Db  2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 595
         ||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 3010
```

RESULT 5
Q81760
```
ID   Q81760           PRELIMINARY;      PRT;   3010 AA.
AC   Q81760;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93359897; PubMed=8394876;
RA   Wang Y., Okamoto H., Tsuda F., Nagayama R., Tao Q.M., Mishiro S.;
RT   "Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus in
RT   Chinese patients with liver disease.";
RL   J. Med. Virol. 40:254-260(1993).
DR   EMBL; D10934; BAA01728.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
```

```
DR   PIR; PQ0254; PQ0254.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;  326857 MW;  EA7D306A4BA2E224 CRC64;

Query Match             95.8%;  Score 3034;  DB 2;  Length 3010;
Best Local Similarity   96.6%;  Pred. No. 1.8e-227;
Matches  571;  Conservative   14;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 2479

Qy       65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            ||||||||||||||||||||||||:|||||||||||||||||||||||:|||:||||||
Db     2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 2539

Qy      125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db     2540 SVWKDLLEDTDTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy      185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVEESIY 244
            ||||||||||||||||||:|||||||||||:||||||||||||||||||||||||:|||||
Db     2600 TLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFPDSTVTENDIRIEESIY 2659

Qy      245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||:||:|||||||:||||:|||||||||||||||||||||||||||||||||
Db     2660 QCCDLAPEAKQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy      305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy      365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            ||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy      425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||:||||||||||||||||||:|||||||||||||||||
Db     2840 ARMILMTHFFSILLAQEQLEKALECQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy      485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db     2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy      545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVIYLLPNR 595
            |||||||||||||||:||||||||||||||||||||||||||||||||||
Db     2960 AASRLDLSGWFVAGYGGGDIYHSLSRARPRWFMLCLLLSVGVIYLLPNR 3010
```

```
RESULT 6
Q02829
ID   Q02829       PRELIMINARY;      PRT;    3010 AA.
AC   Q02829;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HC-J4;
RX   MEDLINE=93391112; PubMed=1325713;
RA   Okamoto H., Kojima M., Okada S., Yoshizawa H., Iizuka H., Tanaka T.,
RA   Muchmore E.E., Peterson D.A., Ito Y., Mishiro S.;
RT   "Genetic drift of hepatitis C virus during an 8.2-year infection in a
RT   chimpanzee: variability and stability.";
RL   Virology 190:894-899(1992).
DR   EMBL; D13558; BAA02755.1; -.
DR   PIR; A61196; A61196.
DR   PIR; A61196; PQ0246.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
```

```
RESULT 7
Q68285
ID   Q68285          PRELIMINARY;       PRT;   3010 AA.
AC   Q68285;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cho J., Park Y., Lee Y., Yang J., Ryu W.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U16362; AAA52748.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q81755; 1DXP.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;   326915 MW;   29B306FC5B8EBC9E CRC64;

Query Match             95.6%;  Score 3027;  DB 2;  Length 3010;
Best Local Similarity   96.8%;  Pred. No. 6.5e-227;
Matches  572; Conservative   9; Mismatches  10; Indels   0; Gaps   0;

Qy        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db      ... SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL ...
```

```
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   CHAIN       730   1006       NS2 protein.
FT   CHAIN      1007   1615       NS3 protein.
FT   CHAIN         1    191       C protein.
FT   CHAIN      1616   2013       NS4 protein.
FT   CHAIN      2014   3010       NS5 protein.
FT   CHAIN       192    383       E protein.
FT   CHAIN       384    729       NS1/E2 protein.
SQ   SEQUENCE   3010 AA;   326959 MW;   93D46526F3EADF8 CRC64;

Query Match             95.7%;  Score 3030;  DB 2;  Length 3010;
Best Local Similarity   96.4%;  Pred. No. 3.8e-227;
Matches  570; Conservative  16; Mismatches  5; Indels   0; Gaps   0;

Qy        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 2479

Qy       65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db     2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIR 2539

Qy      125 SVVKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            ||||:|||||:|| |||||||||||||||||||||||||||||||||||||||||||||
Db     2540 SVVKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy      185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 2659

Qy      245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db     2660 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy      305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPGDPPQPEYDL 364
            |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db     2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRVFTEAMTRYSAPGDPPQPEYDL 2779

Qy      365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db     2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTIPLARAAWETARHTPVNSWLGNIIMYAPALW 2839

Qy      425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2840 ARMILMTHFFSILLAQEQLEQALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy      485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db     2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 2959

Qy      545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            ||||:|||||||||||||||||||||||||||||:|||||||||||||||
Db     2960 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFLLCLLLLSVGVGIYLLPNR 3010
```

DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA; 326925 MW; FE997D54E805142B CRC64;

Query Match   95.6%; Score 3027; DB 2; Length 3010;
Best Local Similarity   96.8%; Pred. No. 6.5e-227;
Matches 572; Conservative   9; Mismatches   10; Indels   0; Gaps   0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVTHIR 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            ||||||||| ::|:|||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTETPISTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            :||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            ::::::::::|:::|:||||||||||||||||||||||||||||||||||||||||||
Db   2660 QCCDLAPEAKLAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLRDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQBQLEKTLDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   2900 EINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 8
Q81825
ID   Q81825   PRELIMINARY;   PRT;   3010 AA.
AC   Q81825;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   MRNA, complete cds.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cho J.-M.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M96362; AAA45721.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q81755; 1DXP.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.

```
SQ    SEQUENCE   3010 AA;   327098 MW;   737BEF31E3C2B28D CRC64;

      Query Match             95.5%;  Score 3026;  DB 2;  Length 3010;
      Best Local Similarity   96.3%;  Pred. No. 7.8e-227;
      Matches   569;  Conservative   14;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||||||||:|||||||||||:|||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYSTTSRSAGLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVBEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||:|||||||||||||:||||||||:|||||||||||||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSARSKFGYGAKDVRSLSSRAVNHIR 2539

Qy    125 SVWKDLLEDTDPIQTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           ||:|||||||||||::|||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTDPIPTTVMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||||||||||||||||||||||||||||:|:||||||||||||||||||||||||||:
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||:||:|||||||||:|||||||||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEAKLAIKSLTERLYIGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           |||||||:|||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db   2720 AACRAAKLRDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||:||||||||||||||||||||:| ||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPATPARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           |||:|||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMVLMTHFFSILIAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||||||||||:||||||||||||||:|||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSVSRARPRWFMWCLLLLSVGVGIYLLPNR 3010
```

```
RESULT 9
Q9J3H0
ID    Q9J3H0                   PRELIMINARY;   PRT;   3010 AA.
AC    Q9J3H0;
DT    01-OCT-2000  (TrEMBLrel. 15, Created)
DT    01-OCT-2000  (TrEMBLrel. 15, Last sequence update)
DT    01-MAR-2004  (TrEMBLrel. 26, Last annotation update)
DE    Polyprotein.
OS    Hepatitis C virus.
OC    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC    Hepacivirus.
OX    NCBI_TaxID=11103;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=MD22;
RA    Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL    Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AF207763; AAF65953.1; -.
DR    PIR; A61196; A61196.
DR    PIR; PQ0246; PQ0246.
DR    PIR; PQ0254; PQ0254.
DR    PIR; PS0329; PS0329.
DR    HSSP; Q8JYS1; 1CWX.
DR    GO; GO:0016021; C:integral to membrane; IEA.
DR    GO; GO:0019028; C:viral capsid; IEA.
DR    GO; GO:0019031; C:viral envelope; IEA.
DR    GO; GO:0005524; F:ATP binding; IEA.
DR    GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR    GO; GO:0003723; F:RNA binding; IEA.
DR    GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR    GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR    GO; GO:0005198; F:structural molecule activity; IEA.
DR    GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR    GO; GO:0006350; P:transcription; IEA.
DR    GO; GO:0019079; P:viral genome replication; IEA.
DR    GO; GO:0019087; P:viral transformation; IEA.
DR    InterPro; IPR000345; CytC_heme_BS.
DR    InterPro; IPR001410; DEAD.
DR    InterPro; IPR011545; DEAD/DEAH_N.
DR    InterPro; IPR002522; HCV_capsid.
DR    InterPro; IPR002521; HCV_core.
DR    InterPro; IPR002519; HCV_env.
DR    InterPro; IPR002531; HCV_NS1.
DR    InterPro; IPR000745; HCV_NS4a.
DR    InterPro; IPR001490; HCV_NS4b.
DR    InterPro; IPR002868; HCV_NS5a.
DR    InterPro; IPR002166; HCV_RdRP.
DR    InterPro; IPR001650; Helicase_C.
DR    InterPro; IPR004109; Peptidase_S29.
DR    InterPro; IPR009003; Pept_Ser_Cys.
DR    InterPro; IPR002518; Pept_U39_HCV_NS2.
DR    InterPro; IPR007095; RNA_pol_DS_PS.
DR    InterPro; IPR007094; RNA_pol_PSvir.
DR    Pfam; PF01543; HCV_capsid; 1.
DR    Pfam; PF01542; HCV_core; 1.
DR    Pfam; PF01539; HCV_env; 1.
DR    Pfam; PF01560; HCV_NS1; 1.
DR    Pfam; PF01538; HCV_NS3; 1.
DR    Pfam; PF02907; HCV_NS3; 1.
DR    Pfam; PF01006; HCV_NS4a; 1.
DR    Pfam; PF01001; HCV_NS4b; 1.
DR    Pfam; PF01506; HCV_NS5a; 1.
DR    Pfam; PF00271; Helicase_C; 1.
DR    Pfam; PF00998; Viral_RdRP; 1.
DR    SMART; SM00487; DEXDc; 1.
DR    PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW    Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW    Polyprotein; Transmembrane.
```

```
RESULT 10
Q9IX6
ID    Q9IX6                    PRELIMINARY;   PRT;   3010 AA.
AC    Q9IX6;
DT    01-MAY-2000  (TrEMBLrel. 13, Created)
DT    01-MAY-2000  (TrEMBLrel. 13, Last sequence update)
DT    01-MAR-2004  (TrEMBLrel. 26, Last annotation update)
DE    Polyprotein.
OS    Hepatitis C virus.
OC    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC    Hepacivirus.
OX    NCBI_TaxID=11103;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=MD8-1;
RX    MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA    Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA    Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT    "Time-related changes in full-length hepatitis C virus and hepatitis
RT    activity.";
RL    Virology 263:244-253(1999).
DR    EMBL; AF165059; AAD56194.1; -.
DR    PIR; A61196; A61196.
DR    PIR; PQ0246; PQ0246.
DR    PIR; PQ0254; PQ0254.
DR    PIR; PQ0804; PQ0804.
```

```
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   MEROPS; S29.002; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_Psvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;  327468 MW;  4613744EC4D4A013 CRC64;

Query Match   95.5%;  Score 3025;  DB 2;  Length 3010;
Best Local Similarity  96.4%;  Pred. No. 9.3e-227;
Matches  570;  Conservative  15;  Mismatches  6;  Indels  0;  Gaps  0;

Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL    64
          ||||||||||||||||||||||||:|| |||||||||:||:||||||||||||||||||||
Db  2420  SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL  2479

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR   124
          ||||||||||||||||||||||||:||||||||||||:|||||||||||||::||:|||||
Db  2480  DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIR  2539

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   184
          |||:||||||:||| |||||||:|||||||||||||||||||||||||||||||||||||
Db  2540  SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  2599

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY   244
          |||||||||||||||||||||||||||||:||||||||||||||||||||:||||:||||
Db  2600  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRTESIY  2659

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNLTCYLKAA   304
          ||||||||||:||:|||||||:||||:|||||||||||||||||||||||||||||||:
Db  2660  QCCDLAPEARKAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNLTCYLKAS  2719
```

```
Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL   364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2720  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  2779

Qy   365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW   424
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  2780  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  2839

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG   484
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db  2840  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  2899

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKTPIP   544
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db  2900  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKTPIP  2959

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMCLLLLSVGVGIYLLPNR   595
          |||||||||||||||||||||||||||||||||| |||||||||||||||
Db  2960  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR  3010
```

```
RESULT 11
Q9QIX5
ID   Q9QIX5        PRELIMINARY;       PRT;    3010 AA.
AC   Q9QIX5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD8-2;
RX   MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA   Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA   Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT   "Time-related changes in full-length hepatitis C virus and hepatitis
RT   activity.";
RL   Virology 263:244-253(1999).
DR   EMBL; AF165060; AAD56195.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
DR   PIR; PQ0254; PQ0254.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   MEROPS; S29.002; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
```

```
AC   P89966;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   RNA for polyprotein, complete cds.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type 1b;
RA   TANAKA T.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type 1b;
RA   Tanaka T.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D89872; BAA14035.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0246; PQ0246.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
     HSSP; Q81755; 1DXP.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_pS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;  327024 MW;  E075BD9CFD8D1261 CRC64;
```

```
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_pS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA;  327295 MW;  8B99F1EBA6A50F56 CRC64;

Query Match             95.5%;  Score 3024;  DB 2;  Length 3010;
Best Local Similarity   96.4%;  Pred. No. 1.1e-226;
Matches  570;  Conservative  14;  Mismatches   7;  Indels   0;  Gaps   0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTPDRLQVL 64
          ||||||||||||||||||||||||:|||||||||||:||:|||||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTPDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          ||||||||||||||||||||||||:|||||||||||:|||||||||||:|||::|||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIR 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          :||:|||||| |||:|||||| ||||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||||||||||||||:||||||||||||||||||||:|||:|||||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRTEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||:||:|||||||:|||:|||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEARKAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 3010
```

```
RESULT 12
P89966
ID   P89966    PRELIMINARY;    PRT;    3010 AA.
```

```
Best Local Similarity  96.44%; Pred. No. 1.6e-226;
Matches 570;  Conservative 12;  Mismatches 9;  Indels 0;  Gaps 0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||::||:|||||||||||||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMVYSTTSRSASLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||||::||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKARLLSIEEACKLTPPHSAKSKFGYGAKDVRSLSSKAVNHIH 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||:||||||:|:|||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            |||||||||||||||||:|||||||||:|||||||||||||||||||||||||||||||||
Db   2600 TLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRTEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETVRHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            ||:||||||||||:||||||:|||||||||||||||||||||||:|||||||
Db   2960 AASQLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 13
Q9DTE0
ID Q9DTE0
AC Q9DTE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Serum;
RA Takahara T., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
DR EMBL; AB049097; BAB18810.1; -.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; 1DXP.
DR MEROPS; S29.002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

```
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003723; F:RNA binding; IEA.
GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F:serine-type peptidase activity; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:viral genome replication; IEA.
GO; GO:0019087; P:viral transformation; IEA.
InterPro; IPR000345; CytC_heme_BS.
InterPro; IPR001410; DEAD.
InterPro; IPR011545; DEAD/DEAH_N.
InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001490; HCV_NS4b.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR004109; Peptidase_S29.
InterPro; IPR009003; Pept_Ser_Cys.
InterPro; IPR002518; Pept_U39_HCV_NS2.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; Helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
SMART; SM00487; DEXDc; 1.
PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
SEQUENCE 3014 AA; 327498 MW; ED9BF4F94BDB6287 CRC64;
```

Query Match   95.4%; Score 3022; DB 2; Length 3014;
Best Local Similarity   96.1%; Pred. No. 1.6e-226;
Matches 568; Conservative 15; Mismatches 8; Indels 0; Gaps 0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||::||:|||||||||||||||||||||||||||||
Db   2424 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMYATTSRSAGLRQKKVTFDRLQVL 2483

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||||::||||||||||
Db   2484 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKYGYGAKDVRSLSSKAVNHIH 2543

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||:||||||:|:||||||||||||||||||||||||||||||||||||||||||||||||
Db   2544 SVWKDLLEDTETPINTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2603

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db   2604 TLPQAVMGSSYGFQYSPQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 2663

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2664 QCCDLAPEAKLAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2723

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2724 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2783

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
```

```
Db    2784 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPIARAAWETARHTPVNSWLGNIIMYAPTLW 2843
           |||||||||||||||||||||| :||||||||||||||||||||||||||||||||||||
Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db    2844 ARMILMTHFFSILLVQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2903
Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db    2904 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2963
Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 595
           ||||||||||||||||||||||||||||||||: ||||||||||||||||
Db    2964 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 3014
```

RESULT 14
009796
ID   009796   PRELIMINARY;   PRT;   3010 AA.
AC   009796;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Genomic RNA, complete cds.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=genotype II/1b;
RX   MEDLINE=97170750; PubMed=9018054;
RA   Sugiyama K., Kato N., Mizutani T., Ikeda M., Tanaka T., Shimotohno K.;
RT   "Genetic analysis of the hepatitis C virus (HCV) genome from HCV-
RT   infected human T cells.";
RL   J. Gen. Virol. 78:329-336(1997).
DR   EMBL; D85516; BAA19625.1; -.
DR   PIR; A61196; A61196.
DR   PIR; PQ0804; PQ0804.
DR   PIR; PS0329; PS0329.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006351; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_NS5a.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.

DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3010 AA; 327266 MW; 4324643B6A1D5CE4 CRC64;

Query Match   95.4%; Score 3021; DB 2; Length 3010;
Best Local Similarity   96.3%; Pred. No. 1.9e-226;
Matches 569; Conservative 14; Mismatches   8; Indels   0; Gaps   0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||| |||||||||||:||:::|||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||| |||||||||||||||| |||||||:||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKYGYGAKDVRNLSSRAVNHIR 2539

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          |||:|||:||:||| ||||||:|||||||||||||||||||||||||||||||||||||||
Db   2540 SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||||:|||
Db   2600 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||:|||||||||||:|||:|||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEAKLAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPBYDL 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPBYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||| :|||||:|||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPIARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||| ||||||||||||:||||||||||||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARGVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 595
          ||||:||||||||||||||||||||||||||||: ||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 3010
```

RESULT 15
039930
ID   039930   PRELIMINARY;   PRT;   591 AA.
AC   039930;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RNA-dependent RNA-polymerase (Fragment).
GN   Name=NS5b;
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;

RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98001363; PubMed=9343198;
RA   Lohmann V., Korner F., Herian U., Bartenschlager R.;
RT   "Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA
RT   polymerase and identification of amino acid sequence motifs essential
RT   for enzymatic activity.";
RL   J. Virol. 71:8416-8428(1997).
DR   EMBL; Z97730; CAB10747.1; -.
DR   PDB; 1NHU; X-ray; A/B=1-570.
DR   PDB; 1NHV; X-ray; A/B=1-570.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   591 AA;   65753 MW;   6FA9C2F970260349 CRC64;

Query Match   95.4%;   Score 3020; DB 2;   Length 591;
Best Local Similarity 96.6%;   Pred. No. 2.5e-227;
Matches 571; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||||||||||||||||||||||||||||||:|||:||:|||:|||||||||||||||||
Db   1 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL   60

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db   61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 120

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy 185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||:|::||||||||||:|:|||||||||||||||||||||||||||||
Db 181 TLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 240

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||||||||||||||:|||||||||||||||||||||:|||||||||||||
Db 241 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||:|||||||||||||||||||||:|||||||||||||||||||||||
Db 301 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRAFTEAMTRYSAPPGDPPQPEYDL 360

Qy 365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db 361 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 420

Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db 421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 480

Qy 485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||:||||||||||:|||||||||||||||||||||
Db 481 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 540

Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||:||:|||| ||||||||||||||||||||||| |||||||||||||||||
Db 541 AASQLDLSSWFVAGYSGGDIYHSLSRARPRWFMWCLLLLSVGVGIYLLPNR 591

Search completed: September 22, 2005, 14:44:15
Job time : 81 secs

This Page Blank (uspto)

```
                 GenCore version 5.1.6
         Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 22, 2005, 14:39:31 ; Search time 27 Seconds
                                              (without alignments)
                              1661.633 Million cell updates/sec

Title:          US-10-712-479-2
Perfect score:  3167
Sequence:       1 MASMSMSYTWTGALITPCAA..........LLSVGVGIYLLPNRHHHHH 601

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:     513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued Patents AA:*
         1: /cgn2_6/ptcdata/1/iaa/5A_COMB.pep:*
         2: /cgn2_6/ptcdata/1/iaa/5B_COMB.pep:*
         3: /cgn2_6/ptcdata/1/iaa/6A_COMB.pep:*
         4: /cgn2_6/ptcdata/1/iaa/6B_COMB.pep:*
         5: /cgn2_6/ptcdata/1/iaa/PCTUS_COMB.pep:*
         6: /cgn2_6/ptcdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	3016	95.2	591	3	US-08-952-981A-1	Sequence 1, Appli
2	3016	95.2	997	1	US-08-324-977-50	Sequence 50, Appl
3	3016	95.2	997	1	US-08-384-616-50	Sequence 50, Appl
4	3016	95.2	997	3	US-08-904-686A-50	Sequence 50, Appl
5	3016	95.2	997	3	US-09-315-850-50	Sequence 50, Appl
6	3016	95.2	2201	2	US-08-952-981A-2	Sequence 2, Appli
7	3016	95.2	2621	2	US-08-324-977-36	Sequence 36, Appl
8	3016	95.2	2621	2	US-08-384-616-36	Sequence 36, Appl
9	3016	95.2	2621	3	US-08-904-686A-36	Sequence 36, Appl
10	3016	95.2	2621	3	US-09-315-850-36	Sequence 36, Appl
11	3016	95.2	3010	1	US-08-324-977-2	Sequence 2, Appli
12	3016	95.2	3010	1	US-08-324-977-14	Sequence 14, Appl
13	3016	95.2	3010	2	US-08-384-616-2	Sequence 2, Appli
14	3016	95.2	3010	2	US-08-384-616-14	Sequence 14, Appl
15	3016	95.2	3010	2	US-08-904-686A-2	Sequence 2, Appli
16	3016	95.2	3010	3	US-08-904-686A-14	Sequence 14, Appl
17	3016	95.2	3010	3	US-09-315-850-2	Sequence 2, Appli
18	3016	95.2	3010	3	US-09-315-850-14	Sequence 14, Appl
19	3015	95.2	3010	3	US-09-014-416-3	Sequence 3, Appli
20	3014	95.2	593	4	US-09-597-877-23	Sequence 23, Appl
21	3011	95.1	621	4	US-09-898-297-1	Sequence 1, Appli
22	3011	95.1	621	4	US-09-995-099-1	Sequence 1, Appli
23	3011	95.1	621	4	US-10-238-282-1	Sequence 1, Appli
24	3011	95.1	2620	1	US-08-324-977-32	Sequence 32, Appl
25	3011	95.1	2620	2	US-08-384-616-32	Sequence 32, Appl
26	3011	95.1	2620	2	US-08-904-686A-32	Sequence 32, Appl
27	3011	95.1	2620	3	US-09-315-850-32	Sequence 32, Appl
28	3005	94.9	1985	4	US-09-539-601-9	Sequence 9, Appli
29	3005	94.9	1985	4	US-09-539-601-12	Sequence 12, Appl
30	3005	94.9	1985	4	US-09-539-601-24	Sequence 24, Appl
31	3005	94.9	2201	4	US-09-539-601-6	Sequence 6, Appli
32	3005	94.9	2201	4	US-09-539-601-15	Sequence 15, Appl
33	3005	94.9	2201	4	US-10-029-907-3	Sequence 3, Appli
34	3005	94.9	3010	4	US-09-539-601-3	Sequence 3, Appli
35	3004	94.9	3010	4	US-09-539-601-27	Sequence 27, Appl
36	3004	94.9	1995	4	US-09-539-601-18	Sequence 18, Appl
37	3004	94.9	3010	4	US-09-539-601-21	Sequence 21, Appl
38	3002	94.8	1985	4	US-09-539-601-30	Sequence 30, Appl
39	3002	94.8	3010	4	US-09-539-601-33	Sequence 33, Appl
40	2993	94.5	591	4	US-09-720-095A-2	Sequence 2, Appli
41	2987	94.3	593	4	US-09-597-877-12	Sequence 12, Appl
42	2927	92.4	590	4	US-09-194-949A-15	Sequence 15, Appl
43	2879	90.9	576	4	US-09-541-990A-1	Sequence 1, Appli
44	2786	88.0	591	3	US-09-208-140-2	Sequence 2, Appli
45	2774	87.6	3011	3	US-09-014-416-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-952-981A-1
; Sequence 1, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
;  APPLICANT: DE FRANCESCO, Raffaele
;  APPLICANT: TOMEI, Licia
;  APPLICANT: BEHRENS, Sven-Erik
;  TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
;  TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
;  TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
;  FILE REFERENCE: IT0002P
;  CURRENT APPLICATION NUMBER: US/08/952,981A
;  CURRENT FILING DATE: 1998-03-23
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 1
;  LENGTH: 591
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-08-952-981A-1

  Query Match             95.2%;  Score 3016;  DB 3;  Length 591;
  Best Local Similarity   96.6%;  Pred. No. 2e-298;
  Matches  571;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||||||||||| ||:||:||||:|||||||||||||
Db      1 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 60

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||||:||:|:||||:
Db     61 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 120

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||| ||:|:||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            |||:|||||||||||||| ||||||||||:||:|||||||||||||||||||||||||||
Db    181 TLPQVVMGSSYGFQYSPCQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 240

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||:|||||||:|||:|||||||||||||||||||||||||||||||||:
Db    241 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 360
```

```
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 997 amino acids
;         TYPE: amino acid
;         TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-08-324-977-50

Query Match            95.2%;  Score 3016;  DB 1;  Length 997;
Best Local Similarity  96.6%;  Pred. No. 4.8e-298;
Matches  571;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||||||||||| ||  |||||| ||||||||||||||
Db  407 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 466

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||: ||||:|
Db  467 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 526

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||  |||  |||||||||||||||||||||||||||||||||||||||||||
Db  527 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 586

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||| ||||||||||| |||||||||||  |  ||||||||||||||||||||||||||||
Db  587 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKNPMGFSYDTRCFDSTVTENDIRVEESIY 646

Qy  245 QCCDLAPEARQAIRSLITERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||:|| |||||  ||  |||||||||||||||||||||||||||||||||:
Db  647 QCCDLAPEARQAIKSLTERLYGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 706

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  707 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 766

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  767 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 826

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  827 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 886

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||:||||||||| ||||||||||:|||||||||
Db  887 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 946

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  947 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 997


RESULT 3
US-08-384-616-50
; Sequence 50, Application US/0838416
; Patent No. 5847101
; GENERAL INFORMATION:
;    APPLICANT: OKAYAMA, Hiroto
;    APPLICANT: FUKE, Isao
;    APPLICANT: MORI, Chisato
;    APPLICANT: TAKAMIZAWA, Akahisa
;    APPLICANT: YOSHIDA, Iwao
;    TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES: 50
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;       ADDRESSEE: Naughton
;       STREET: 1725 K St. N.W. Suite 1000
;       CITY: Washington
;       STATE: D.C.
;       COUNTRY: U.S.A.
;       ZIP: 20006
```

```
Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  361 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 420

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          |||||||||||||||||||||||||||||:|||||||||| ||||||||||:|||||||||
Db  481 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 540

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 591


RESULT 2
US-08-324-977-50
; Sequence 50, Application US/0832497
; Patent No. 5747339
; GENERAL INFORMATION:
;    APPLICANT: OKAYAMA, Hiroto
;    APPLICANT: FUKE, Isao
;    APPLICANT: MORI, Chisato
;    APPLICANT: TAKAMIZAWA, Akahisa
;    APPLICANT: YOSHIDA, Iwao
;    TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;    TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;    NUMBER OF SEQUENCES: 50
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;       ADDRESSEE: Naughton
;       STREET: 1725 K St. N.W. Suite 1000
;       CITY: Washington
;       STATE: D.C.
;       COUNTRY: U.S.A.
;       ZIP: 20006
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;       SOFTWARE: ASCII
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/324,977
;       FILING DATE: 18-OCT-1994
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 08/099,706
;       FILING DATE: 30-JUL-1993
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 07/769,996
;       FILING DATE: 02-OCT-1991
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 07/635,451
;       FILING DATE: 28-DEC-1990
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 2-167466
;       FILING DATE: 25-JUN-1990
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 2-305605
;       FILING DATE: 09-NOV-1990
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: JP 2-230921
;       FILING DATE: 31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Stevens-Smith, Theresa M.
;       REGISTRATION NUMBER: 36,281
;       REFERENCE/DOCKET NUMBER: 900703D
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (202) 659-2930
;       TELEFAX: (202) 887-0357
;       TELEX: 440142
; INFORMATION FOR SEQ ID NO: 50:
```

```
Db  827 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 886
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||||||||||||||||||||||||||::||||||||| :||||||:|||||||||
Db  887 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 946
Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            |||||||||||||||||||||||| |||||||||||||||||||||||
Db  947 AASRLDLSGWFVAGYSGGDIYHSLSLSRARPRWFMLCLLLLSVGVGIYLLPNR 997
```

RESULT 4
US-08-904-686A-50
; Sequence 50, Application US/08904686A
; Patent No. 5998130
GENERAL INFORMATION:
  APPLICANT: OKAYAMA, Hiroto
  APPLICANT: FUKE, Isao
  APPLICANT: MORI, Chisato
  APPLICANT: TAKAMIZAWA, Akahisa
  APPLICANT: YOSHIDA, Iwao
  TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
  TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
  NUMBER OF SEQUENCES: 50
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
    ADDRESSEE: Naughton
    STREET: 1725 K St. N.W. Suite 1000
    CITY: Washington
    STATE: D.C.
    COUNTRY: U.S.A.
    ZIP: 20006
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
    SOFTWARE: ASCII
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/904,686A
    FILING DATE: 01-AUG-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/324,977
    FILING DATE: 18-OCT-1994
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: JP 2-167466
    FILING DATE: 25-JUN-1990
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: JP 2-230921
    FILING DATE: 31-AUG-1990
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: JP 2-305605
    FILING DATE: 09-NOV-1990
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/099,706
    FILING DATE: 30-JUL-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/769,996
    FILING DATE: 02-OCT-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/635,451
    FILING DATE: 28-DEC-1990
  ATTORNEY/AGENT INFORMATION:
    NAME: McLeland, Le-Nhung
    REGISTRATION NUMBER: 31,541
    REFERENCE/DOCKET NUMBER: 900703G
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (202) 659-2930
    TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 50:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 997 amino acids
    TYPE: amino acid
    TOPOLOGY: linear

COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
  SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/384,616
  FILING DATE:
  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/769,996
  FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: JP 2-167466
  FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: JP 2-230921
  FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: JP 2-305605
  FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/635,451
  FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
  NAME: Stevens-Smith, Theresa M.
  REGISTRATION NUMBER: 36,281
  REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 659-2930
  TELEFAX: (202) 887-0357
  TELEX: 440142
INFORMATION FOR SEQ ID NO: 50:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 997 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-384-616-50

Query Match   95.2%; Score 3016; DB 2; Length 997;
Best Local Similarity   96.6%; Pred. No. 4.8e-298;
Matches   571; Conservative   11; Mismatches   9; Indels   0; Gaps   0;

```
Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL 64
          ||||||||||||||||||||||||||||||||||||||:||:|||||:||||||||||
Db  407 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 466

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||:
Db  467 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPFGYGAKDVRNLSSKAVNHIH 526

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||| |||:||||||||||||||||||||||||||||||||||||||||||||
Db  527 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 586

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||||.||||||||||||.|||||||||||.||..|||||||||||||||||||||||||||
Db  587 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 646

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||:||||||||:||:||||||||||||||||||||||||||||||||||:
Db  647 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 706

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  707 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 766

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  767 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 826

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;   MOLECULE TYPE:  protein
US-08-904-686A-50

Query Match             95.2%;  Score 3016;  DB 2;  Length 997;
Best Local Similarity   96.6%;  Pred. No. 4.8e-298;
Matches   571;  Conservative  11;  Mismatches   9;  Indels  9;  Gaps  0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL  64
            ||||||||||||||||||||||||||||||:|:||:||||||||||:|||||||||||
Db    407 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 466

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||:||||:|||||:||
Db    467 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 526

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db    527 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 586

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||||||||||||:||||||||||||.||:||||||||||||||||||||||||||
Db    587 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 646

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    647 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 706

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    707 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 766

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||:||||:||||||||||||||||||||||||||||
Db    767 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 826

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    827 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 886

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    887 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 946

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            |||||||||||||||||||||||||||||||:|||||||||||||||||||
Db    947 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 997


RESULT 5
US-09-315-850-50
; Sequence 50, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
;  APPLICANT:  OKAYAMA, Hiroto
;  APPLICANT:  FUKE, Isao
;  APPLICANT:  MORI, Chisato
;  APPLICANT:  TAKAMIZAWA, Akahisa
;  APPLICANT:  YOSHIDA, Iwao
;  TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOMIC
;  TITLE OF INVENTION:  CDNA AND ANTIGEN POLYPEPTIDE
;  NUMBER OF SEQUENCES:  50
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Armstrong, Westerman, Hattori, McLeland, &
;   ADDRESSEE:  Naughton
;   STREET:  1725 K St. N.W. Suite 1000
;   CITY:  Washington
;   STATE:  D.C.
;   COUNTRY:  U.S.A.
;   ZIP:  20006
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Diskette, 3.5 in, 1.44Mb
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS, Version 5.0
```

```
;      SOFTWARE:  ASCII
;   CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/09/315,850
;    FILING DATE:
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US/08/904,686
;    FILING DATE:  01-AUG-1997
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 08/324,977
;    FILING DATE:  18-OCT-1994
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  JP 2-167466
;    FILING DATE:  25-JUN-1990
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  JP 2-230921
;    FILING DATE:  31-AUG-1990
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  JP 2-305605
;    FILING DATE:  09-NOV-1990
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 08/099,706
;    FILING DATE:  30-JUL-1993
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 07/769,996
;    FILING DATE:  02-OCT-1991
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US 07/635,451
;    FILING DATE:  28-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;    NAME:  McLeland, Le-Nhung
;    REGISTRATION NUMBER:  31,541
;    REFERENCE/DOCKET NUMBER:  900703G
;   TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  (202) 659-2930
;    TELEFAX:  (202) 887-0357
; INFORMATION FOR SEQ ID NO:  50:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH:  997 amino acids
;    TYPE:  amino acid
;    TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-09-315-850-50

Query Match             95.2%;  Score 3016;  DB 3;  Length 997;
Best Local Similarity   96.6%;  Pred. No. 4.8e-298;
Matches   571;  Conservative  11;  Mismatches   9;  Indels  9;  Gaps  0;

Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVL  64
            ||||||||||||||||||||||||||||||:|:||:||||||||||:|||||||||||
Db    407 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 466

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||||:||||:|||||:||
Db    467 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 526

Qy    125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db    527 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 586

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            ||||||||||||||||:||||||||||||.||:||||||||||||||||||||||||||
Db    587 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 646

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    647 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 706

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    707 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 766

Qy    365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||:||||:||||||||||||||||||||||||||||
Db    767 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 826
```

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   827 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 886

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||:||||||||||::|||||||||:|||||||||
Db   887 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 946

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   947 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 997

RESULT 6
US-08-952-981A-2
; Sequence 2, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
;   APPLICANT: DE FRANCESCO, Raffaele
;   APPLICANT: TOMEI, Licia
;   APPLICANT: BEHRENS, Sven-Brik
;   TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
;   TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
;   TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
;   FILE REFERENCE: IT0002P
;   CURRENT APPLICATION NUMBER: US/08/952,981A
;   CURRENT FILING DATE: 1998-03-23
;   NUMBER OF SEQ ID NOS: 14
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 2
;   LENGTH: 2201
;   TYPE: PRT
;   ORGANISM: cDNA clone pCD (38-9.4)
US-08-952-981A-2

  Query Match   95.2%; Score 3016; DB 3; Length 2201;
  Best Local Similarity   96.6%; Pred. No. 1.8e-297;
  Matches   571; Conservative   11; Mismatches   9; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           ||||||||||||||||||||||||||||||||||:||:||||||:||||||||||||||
Db   1611 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 1670

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||||||||||||||||||||||||||||||||:|||:||||||:
Db   1671 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 1730

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1731 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 1790

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||.||||||||||||.:|||||||||.||.:|||||||||||||||||||||||||||
Db   1791 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 1850

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           ||||||||||||||:||||||:|||:|||||||||||||||||||||||||||||||||:
Db   1851 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 1910

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 1970

Qy   365 ELITSCSSNVSVAHDASGKRVYFLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db   1971 ELITSCSSNVSVAHDASGKRVYFLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2030

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2090

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||:||||||||||::|||||||||:|||||||||
Db   2091 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2150

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2151 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 2201

RESULT 7
US-08-324-977-36
; Sequence 36, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
;   APPLICANT: OKAYAMA, Hiroto
;   APPLICANT: FUKE, Isao
;   APPLICANT: MORI, Chisato
;   APPLICANT: TAKAMIZAWA, Akahisa
;   APPLICANT: YOSHIDA, Iwao
;   TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;   TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;   NUMBER OF SEQUENCES: 50
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;   ADDRESSEE: Naughton
;   STREET: 1725 K St. N.W. Suite 1000
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20006
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;   SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/324,977
;   FILING DATE: 18-OCT-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-167466
;   FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-230921
;   FILING DATE: 31-AUG-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-305605
;   FILING DATE: 09-NOV-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/099,706
;   FILING DATE: 30-JUL-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/769,996
;   FILING DATE: 02-OCT-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/635,451
;   FILING DATE: 28-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Stevens-Smith, Theresa M.
;   REGISTRATION NUMBER: 36,281
;   REFERENCE/DOCKET NUMBER: 900703D
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 659-2930
;   TELEFAX: (202) 887-0357
;   TELEX: 440142
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2621 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-324-977-36

  Query Match   95.2%; Score 3016; DB 1; Length 2621;
  Best Local Similarity   96.6%; Pred. No. 2.4e-297;

PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/769,996
  FILING DATE: 02-OCT-1991
  APPLICATION NUMBER: JP 2-167466
  FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: JP 2-230921
  FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: JP 2-305605
  FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/635,451
  FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
  NAME: Stevens-Smith, Theresa M.
  REGISTRATION NUMBER: 36,281
  REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 659-2930
  TELEFAX: (202) 887-0357
  TELEX: 440142
INFORMATION FOR SEQ ID NO: 36:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 2621 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein

US-08-384-616-36

```
Query Match          95.2%;    Score 3016;  DB 2;  Length 2621;
Best Local Similarity 96.6%;   Pred. No. 2.4e-297;
Matches  571; Conservative  11; Mismatches  9; Indels  0; Gaps  0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
         |||||||||||||||||||||||||||||||||| :|| ::||||| ||||||||||||||
Db  2031 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2090

Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
         ||||||||||||||||||||||||||||||||||||||||||||||||:||::|||||:
Db  2091 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2150

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
         |||||||||| ::|||||||||||||||||||||||||||||||||||||||||||||||
Db  2151 SVWKDLLEDTVPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2210

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
         |||| ||||||||||||  |||||||||| || ||||||||||||||||||||||||||||
Db  2211 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2270

Qy   245 QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
         |||||||||||||:|||||||| ||| :||||||||||||||||||||||||||||||||:
Db  2271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2330

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2390

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  2391 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2450

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2451 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2510

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
         ||||||||||||||||||||||||||||||:|||||||| ||||||||||:|||||||||
Db  2511 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2570

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2571 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 2621
```

RESULT 8
US-08-384-616-36
; Sequence 36, Application US/08384616
; Patent No. 5847101
GENERAL INFORMATION:
  APPLICANT: OKAYAMA, Hiroto
  APPLICANT: FUKE, Isao
  APPLICANT: MORI, Chisato
  APPLICANT: TAKAMIZAWA, Akahisa
  APPLICANT: YOSHIDA, Iwao
  TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
  TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
  NUMBER OF SEQUENCES: 50
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
  ADDRESSEE: Naughton
  STREET: 1725 K St. N.W. Suite 1000
  CITY: Washington
  STATE: D.C.
  COUNTRY: U.S.A.
  ZIP: 20006
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
  SOFTWARE: ASCII
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/384,616
  FILING DATE:
  CLASSIFICATION: 424

```
Db  2031 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2090
Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
Db  2091 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPHSAKSKFGYGAKDVRNLSSKAVNHIH 2150
Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Db  2151 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2210
Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVEESIY 244
Db  2211 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFPDSTVTENDIRVEESIY 2270
Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
Db  2271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2330
Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
Db  2331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2390
Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
Db  2391 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2450
Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
Db  2451 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2510
Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
Db  2511 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2570
Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
Db  2571 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 2621


RESULT 10
US-09-315-850-36
; Sequence 36, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
    APPLICANT: OKAYAMA, Hiroto
    APPLICANT: FUKE, Isao
    APPLICANT: MORI, Chisato
    APPLICANT: TAKAMIZAWA, Akahisa
    APPLICANT: YOSHIDA, Iwao
    TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
    TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
    NUMBER OF SEQUENCES: 50
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
      ADDRESSEE: Naughton
      STREET: 1725 K St. N.W. Suite 1000
      CITY: Washington
      STATE: D.C.
      COUNTRY: U.S.A.
      ZIP: 20006
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
      SOFTWARE: ASCII
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/315,850
      FILING DATE:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/904,686
      FILING DATE: 01-AUG-1997
      APPLICATION NUMBER: US 08/324,977
      FILING DATE: 18-OCT-1994
```

```
RESULT 9
US-08-904-686A-36
; Sequence 36, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
    APPLICANT: OKAYAMA, Hiroto
    APPLICANT: FUKE, Isao
    APPLICANT: MORI, Chisato
    APPLICANT: TAKAMIZAWA, Akahisa
    APPLICANT: YOSHIDA, Iwao
    TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
    TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
    NUMBER OF SEQUENCES: 50
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
      ADDRESSEE: Naughton
      STREET: 1725 K St. N.W. Suite 1000
      CITY: Washington
      STATE: D.C.
      COUNTRY: U.S.A.
      ZIP: 20006
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
      SOFTWARE: ASCII
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/904,686A
      FILING DATE: 01-AUG-1997
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/324,977
      FILING DATE: 18-OCT-1994
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 2-167466
      FILING DATE: 25-JUN-1990
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 2-230921
      FILING DATE: 31-AUG-1990
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 2-230922
      FILING DATE: 31-AUG-1990
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 2-305605
      FILING DATE: 09-NOV-1990
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/099,706
      FILING DATE: 30-JUL-1993
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/769,996
      FILING DATE: 02-OCT-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/635,451
      FILING DATE: 28-DEC-1990
    ATTORNEY/AGENT INFORMATION:
      NAME: McLeland, Le-Nhung
      REGISTRATION NUMBER: 31,541
      REFERENCE/DOCKET NUMBER: 900703G
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (202) 659-2930
      TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 36:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2621 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
US-08-904-686A-36

Query Match             95.2%;  Score 3016;  DB 2;  Length 2621;
  Best Local Similarity   96.6%;  Pred. No. 2.4e-297;
  Matches  571;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYSTTSRSASLRQKKVTFDRLQVL 64
```

```
Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2571  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  2621
```

RESULT 11
US-08-324-977-2
; Sequence 2, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
;   APPLICANT: OKAYAMA, Hiroto
;   APPLICANT: FUKE, Isao
;   APPLICANT: MORI, Chisato
;   APPLICANT: TAKAMIZAWA, Akahisa
;   APPLICANT: YOSHIDA, Iwao
;   TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;   TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;   NUMBER OF SEQUENCES: 50
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;   ADDRESSEE: Naughton
;   STREET: 1725 K St. N.W. Suite 1000
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20006
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;   SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/324,977
;   FILING DATE: 18-OCT-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-167466
;   FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-230921
;   FILING DATE: 31-AUG-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-305605
;   FILING DATE: 09-NOV-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/099,706
;   FILING DATE: 30-JUL-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/769,996
;   FILING DATE: 02-OCT-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/635,451
;   FILING DATE: 28-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Stevens-Smith, Theresa M.
;   REGISTRATION NUMBER: 36,281
;   REFERENCE/DOCKET NUMBER: 900703D
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 659-2930
;   TELEFAX: (202) 887-0357
;   TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3010 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-324-977-2

Query Match   95.2%;   Score 3016;   DB 1;   Length 3010;
Best Local Similarity   96.6%;   Pred. No. 3e-297;
Matches   571;   Conservative   11;   Mismatches   9;   Indels   0;   Gaps   0;

```
Qy   5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
        |||||||||||||||||||||||||||||||||||||||||| :|||||||||||||||
Db   5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
```

PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-167466
;   FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-230921
;   FILING DATE: 31-AUG-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-305605
;   FILING DATE: 09-NOV-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/099,706
;   FILING DATE: 30-JUL-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/769,996
;   FILING DATE: 02-OCT-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/635,451
;   FILING DATE: 28-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: McLeland, Le-Nhung
;   REGISTRATION NUMBER: 31,541
;   REFERENCE/DOCKET NUMBER: 900703G
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 659-2930
;   TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2621 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-315-850-36

Query Match   95.2%;   Score 3016;   DB 3;   Length 2621;
Best Local Similarity   96.6%;   Pred. No. 2.4e-297;
Matches   571;   Conservative   11;   Mismatches   9;   Indels   0;   Gaps   0;

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL     64
          ||||||||||||||||||||||||||||||||||||||| :|||: ||||||||||||||
Db  2031  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL  2090

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPGYGAKDVRSLLSSRAVNHIR   124
          ||||||||||||||||||||||||||||||||||||||| :||||:||||||||||||||
Db  2091  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPGYGAKDVRNLSSKAVNHIH   2150

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   184
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2151  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   2210

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVEESIY   244
          |||| |||||||||||||||||||||||||||| :|||||||||||||||||||||||||
Db  2211  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFPDSTVTENDIRVEESIY   2270

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA   304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2271  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS   2330

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL   364
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2331  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL   2390

Qy   365  ELITSCSSNVSVAHDASGKRVVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW   424
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  2391  ELITSCSSNVSVAHDASGKRVVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW   2450

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG   484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2451  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG   2510

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP   544
          ||||||||||||||||||||||||||||||||||||| ||||||||||:||||||||||
Db  2511  EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP   2570
```

APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-324-977-14

Query Match   95.2%;   Score 3016; DB 1; Length 3010;
Best Local Similarity   96.6%;   Pred. No. 3e-297;
Matches   571; Conservative   11; Mismatches   9; Indels   0; Gaps   0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||||||||||||||| :||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIH 124
          |||||||||||||||||||||||||||||||||||||||:||:|||||||||||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

Qy    125 SVWKDLLEDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          |||||||||:|| ||||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTVPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy    185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||| |||||||||||| ||||||||||:: |||||||||||||||||||||||||
Db   2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy    245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||:|||:||||||||||| || ::|||||||||||||||||||||||||
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy    305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy    365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy    425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy    485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||:|||||||||:||||||||:||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy    545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 12
US-08-324-977-14
; Sequence 14, Application US/08324977
; Patent No. 5747339
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKE, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESSEE: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:

RESULT 13
US-08-384-616-2
; Sequence 2, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
;   APPLICANT: OKAYAMA, Hiroto
;   APPLICANT: FUKE, Isao
;   APPLICANT: MORI, Chisato
;   APPLICANT: TAKAMIZAWA, Akahisa
;   APPLICANT: YOSHIDA, Iwao
;   TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;   TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;   NUMBER OF SEQUENCES: 50
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;   ADDRESSEE: Naughton
;   STREET: 1725 K St. N.W. Suite 1000
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20006
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;   SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/384,616
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/769,996
;   FILING DATE: 02-OCT-1991
;   APPLICATION NUMBER: JP 2-167466
;   FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-230921
;   FILING DATE: 31-AUG-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-305605
;   FILING DATE: 09-NOV-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/635,451
;   FILING DATE: 28-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Stevens-Smith, Theresa M.
;   REGISTRATION NUMBER: 36,281
;   REFERENCE/DOCKET NUMBER: 900703B
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 659-2930
;   TELEFAX: (202) 887-0357
;   TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3010 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-384-616-2

Query Match   95.2%; Score 3016; DB 2; Length 3010;
Best Local Similarity   96.6%; Pred. No. 3e-297;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||||||||||||||:|: ||||||||||||||||||||
Db   2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMYATTSRSAGLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||:||:||||||||:||||||
Db   2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

RESULT 14
US-08-384-616-14
; Sequence 14, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
;   APPLICANT: OKAYAMA, Hiroto
;   APPLICANT: FUKE, Isao
;   APPLICANT: MORI, Chisato
;   APPLICANT: TAKAMIZAWA, Akahisa
;   APPLICANT: YOSHIDA, Iwao
;   TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
;   TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE
;   NUMBER OF SEQUENCES: 50
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
;   ADDRESSEE: Naughton
;   STREET: 1725 K St. N.W. Suite 1000
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20006
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
;   SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/384,616
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/769,996
;   FILING DATE: 02-OCT-1991
;   APPLICATION NUMBER: JP 2-167466
;   FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 2-230921
;   FILING DATE: 31-AUG-1990
;   PRIOR APPLICATION DATA:

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            |||| ||||||||||||| :|||||||||||||||||||||||||||||||||||||||||
Db   2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            |||||||||||||:|:||||||||:|||||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||||||||||||||||||||||||||||||||||||| :|||||||||:|||:|||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010

APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-384-616-14

```
Query Match      95.2%;      Score 3016;  DB 2;  Length 3010;
Best Local Similarity 96.6%;  Pred. No. 3e-297;
Matches 571;  Conservative 11;  Mismatches 9;  Indels 0;  Gaps 0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          |||||||||||||||||||||||||||||||||:|||||||||| ||:||||||||||||
Db 2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          |||||||||||||||||||||||||||||||||||||||||||||:|||::|||:|||||
Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||||:|:|||||||||||||||||||||||||||||||||||||||||||||
Db 2540 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          ||||||||||||||||||||||||||||:|||||||||||||||||||||||:|||||
Db 2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db 2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
          |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db 2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
          |||||||||||||||||||||||||||||:||||||||||:|||||||||:||||||||
Db 2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010
```

RESULT 15
US-08-904-686A-2
; Sequence 2, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKE, Isao

APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESSEE: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLeland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-904-686A-2

```
Query Match      95.2%;      Score 3016;  DB 2;  Length 3010;
Best Local Similarity 96.6%;  Pred. No. 3e-297;
Matches 571;  Conservative 11;  Mismatches 9;  Indels 0;  Gaps 0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          |||||||||||||||||||||||||||||||||:|||||||||| ||:||||||||||||
Db 2420 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 2479

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          |||||||||||||||||||||||||||||||||||||||||||||:|||::|||:|||||
Db 2480 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 2539

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||||:|:|||||||||||||||||||||||||||||||||||||||||||||
Db 2540 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 2599
```

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           |||||||||||||||||||||: |||: ||||||||||||||||||||||||||||||||
Db   2600 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 2659

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||||||||||: |||: |||: ||||||||||||||||||||||||||||||||:
Db   2660 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 2719

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 2779

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 2839

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 2899

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||||: ||||||||||||||: ||||||||||
Db   2900 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 2959

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR 595
           ||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   2960 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 3010

Search completed: September 22, 2005, 14:47:44
Job time : 30 secs

```
          GenCore version 5.1.6
     Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:  September 22, 2005, 14:43:14 ; Search time 76 Seconds
                                        (without alignments)
                                        3218.609 Million cell updates/sec

Title:           US-10-712-479-2
Perfect score:   3167
Sequence:        1 MASMSMSYTWTGALITPCAA.........LLSVGVGIYLLPNRHHHHH 601

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1826521 seqs, 407012169 residues

Total number of hits satisfying chosen parameters:   1826521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published Applications AA:*
      1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
      2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
      3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
      4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
      5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
      6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
      7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
      8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
      9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
     10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
     11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
     12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
     13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
     14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
     15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
     16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
     17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
     18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
     19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
     20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
     21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
     22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	3167	100.0	601	16	US-10-712-479-2	Sequence 2, Appli
2	3034.5	95.8	580	16	US-10-712-479-4	Sequence 4, Appli
3	3016	95.2	591	13	US-10-085-476-1	Sequence 1, Appli
4	3016	95.2	2201	13	US-10-085-476-2	Sequence 2, Appli
5	3014	95.2	593	14	US-10-241-872-23	Sequence 23, Appl
6	3012	95.1	1985	16	US-10-492-178-6	Sequence 6, Appli
7	3011	95.1	591	16	US-10-471-164-5	Sequence 5, Appli
8	3011	95.1	621	9	US-09-898-297-1	Sequence 1, Appli
9	3011	95.1	621	14	US-10-198-680A-1	Sequence 1, Appli
10	3011	95.1	621	15	US-10-238-282-1	Sequence 1, Appli
11	3011	95.1	621	15	US-10-198-259A-1	Sequence 1, Appli
12	3011	95.1	621	15	US-10-198-384A-1	Sequence 1, Appli
13	3011	95.1	621	16	US-10-471-164-1	Sequence 1, Appli
14	3011	95.1	621	16	US-10-851-710-1	Sequence 1, Appli
15	3011	95.1	627	9	US-09-838-386-12	Sequence 12, Appl
16	3011	95.1	627	16	US-10-471-164-6	Sequence 6, Appli
17	3006	94.9	2865	9	US-09-742-659-6	Sequence 42, Appl
18	3005	94.9	1985	14	US-10-259-275-42	Sequence 42, Appl
19	3005	94.9	1985	16	US-10-639-150-2	Sequence 2, Appli
20	3005	94.9	1985	20	US-11-006-313-42	Sequence 3, Appli
21	3005	94.9	2201	13	US-10-029-907-3	Sequence 3, Appli
22	3005	94.9	2201	14	US-10-309-561-3	Sequence 3, Appli
23	3005	94.9	2201	16	US-10-789-355-3	Sequence 3, Appli
24	3005	94.9	2201	16	US-10-686-835-3	Sequence 1, Appli
25	3010	94.9	3010	15	US-10-467-000-1	Sequence 12, Appl
26	2991	94.4	1985	16	US-10-492-178-1	Sequence 40, Appl
27	2987	94.3	593	14	US-10-241-872-12	Sequence 34, Appl
28	2978	94.0	2985	14	US-10-259-275-40	Sequence 15, Appli
29	2978	94.0	2985	20	US-11-006-313-40	Sequence 15, Appli
30	2972	93.8	3010	16	US-10-333-449A-34	Sequence 3, Appli
31	2937.5	92.8	576	17	US-10-842-046-1	Sequence 2, Appli
32	2927	92.4	590	10	US-09-194-949-15	Sequence 4, Appli
33	2927	92.4	590	17	US-10-664-391-15	Sequence 2, Appli
34	2925.5	92.4	578	14	US-10-211-455-1	Sequence 2, Appli
35	2925.5	92.4	578	16	US-10-471-164-3	Sequence 1, Appli
36	2904	91.7	591	14	US-10-211-455-2	Sequence 2, Appli
37	2904	91.7	591	14	US-10-211-455-4	Sequence 2, Appli
38	2904	91.7	591	16	US-10-471-164-2	Sequence 1, Appli
39	2879	90.9	576	14	US-10-170-131-1	Sequence 2, Appli
40	2786	88.0	591	9	US-09-733-183A-2	Sequence 2, Appli
41	2786	88.0	591	14	US-10-342-372-2	Sequence 2, Appli
42	2776	87.7	3011	15	US-10-296-734-406	Sequence 406, App
43	2774	87.6	3011	9	US-09-742-659-4	Sequence 4, Appli
44	2774	87.6	3011	9	US-09-952-572-9	Sequence 9, Appli
45	2774	87.6	3011	9	US-09-747-419-20	Sequence 20, Appl

ALIGNMENTS

```
RESULT 1
US-10-712-479-2
; Sequence 2, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
;  APPLICANT: Yagi, Yoshihiko
;  APPLICANT: Sheets, Michael P.
;  APPLICANT: Wells, Peter A.
;  APPLICANT: Shelly, John A.
;  APPLICANT: Poorman, Roger A.
;  APPLICANT: Epps, Dennis E.
;  TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Polym
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 02-333-A
;  CURRENT APPLICATION NUMBER: US/10/712,479
;  CURRENT FILING DATE: 2003-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;  LENGTH: 601
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-10-712-479-2

  Query Match             100.0%;  Score 3167;  DB 16;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 4e-280;
  Matches  601;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60

Qy         61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120

Db  121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180

Db  181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVE 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVE 240

Db  241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300

Db  301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360

Db  361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420

Db  421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480

Db  481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540

Db  541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNRHHHHH 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNRHHHHH 600

Db  601 H 601
          |
Qy  601 H 601
```

```
Qy   61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV 120

Qy  121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY 180

Qy  181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVE 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFPDSTVTENDIRVE 240

Qy  241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY 300

Qy  301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP 360

Qy  361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA 420

Qy  421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS 480

Qy  481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL 540

Qy  541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARP----------------------- 579
          ||||||||||||||||||||||||||||||||||                        RHHHHH
Db  541 TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARP

Db  580 H 580
```

```
RESULT 2
US-10-712-479-4
; Sequence 4, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
;  APPLICANT: Yagi, Yoshihiko
;  APPLICANT: Sheets, Michael P.
;  APPLICANT: Wells, Peter A.
;  APPLICANT: Shelly, John A.
;  APPLICANT: Poorman, Roger A.
;  APPLICANT: Epps, Dennis E.
;  TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Poly
;  FILE REFERENCE: 02-333-A
;  CURRENT APPLICATION NUMBER: US/10/712,479
;  CURRENT FILING DATE: 2003-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 4
;  LENGTH: 580
;  TYPE: PRT
;  ORGANISM: Artificial
;  FEATURE:
;  OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NS5B)
US-10-712-479-4

  Query Match             95.8%;  Score 3034.5;  DB 16;  Length 580;
  Best Local Similarity   96.5%;  Pred. No. 4.9e-268;
  Matches   580;  Conservative   0;  Mismatches   0;  Indels  21;  Gaps   1;

Qy    1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR 60
```

```
RESULT 3
US-10-085-476-1
; Sequence 1, Application US/10085476
; Publication No. US20020164722A1
; GENERAL INFORMATION:
;  APPLICANT: De Francesco, Raffaele
;  APPLICANT: Tomei, Licia
;  APPLICANT: Behrens, Sven-Erik
;  TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE
;  TITLE OF INVENTION: RNA-DEPENDENT RNA POLYMERASE AND TERMINAL NUCLEOTIDYL
;  TITLE OF INVENTION: TRANSFERASE ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
;  FILE REFERENCE: IT0002PCA
;  CURRENT APPLICATION NUMBER: US/10/085,476
;  CURRENT FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 08/952,981
;  PRIOR FILING DATE: 1998-03-23
;  PRIOR APPLICATION NUMBER: PCT/IT96/00106
;  PRIOR FILING DATE: 1996-05-24
;  PRIOR APPLICATION NUMBER: RM95A000343
;  PRIOR FILING DATE: 1995-05-25
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastpSEQ for Windows Version Version 4.0
;  SEQ ID NO 1
;  LENGTH: 591
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-10-085-476-1

  Query Match             95.2%;  Score 3016;  DB 13;  Length 591;
  Best Local Similarity   96.6%;  Pred. No. 2.5e-266;
  Matches   571;  Conservative   11;  Mismatches   9;  Indels   0;  Gaps   0;
```

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||||||||||||||||||||||||||||||||| :||:|||||| :||||||||||||
Db     1  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL   60

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          |||||||||||||||||||||||||||||||||||||||||||||||| :|||:|||||:
Db    61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH  120

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db   121  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          |||| |||||||||||| |||||||||| :|| ||||||||||||||||||||||||||
Db   181  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY  240

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          |||||||||||||:||||||:|||::||||||||||||||||||||||||||||||||:|
Db   241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  360

Qy   365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   361  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  420

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKTPIP   544
          ||||||||||||||||||||||||||||:|||||||||| ||||||||||:|||||||||
Db   481  EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKTPIP   540

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR   595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR   591
```

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL    64
          ||||||||||||||||||||||||||||||||||| :||:|||||| :||||||||||||
Db  1611  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL  1670

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR   124
          |||||||||||||||||||||||||||||||||||||||||||||||| :|||:|||||:
Db  1671  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH  1730

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   184
          |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db  1731  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  1790

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY   244
          |||| |||||||||||| |||||||||| :|| ||||||||||||||||||||||||||
Db  1791  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY  1850

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA   304
          |||||||||||||:||||||:|||::||||||||||||||||||||||||||||||||:|
Db  1851  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  1910

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL   364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1911  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  1970

Qy   365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW   424
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1971  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  2030

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG   484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2031  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  2090

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKTPIP    544
          ||||||||||||||||||||||||||||:|||||||||| ||||||||||:|||||||||
Db  2091  EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKTPIP   2150

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR   595
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2151  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR  2201
```

Query Match 95.2%; Score 3016; DB 13; Length 2201;
Best Local Similarity 96.6%; Pred. No. 1.7e-265;
Matches 571; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Query Match 95.2%; Score 3014; DB 14; Length 593;
Best Local Similarity 96.3%; Pred. No. 3.8e-266;
Matches 569; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||||||||||||||||||||||||||||||||| :||:|||||| :||||||||||||
```

```
Query Match          95.1%;   Score 3012; DB 16;  Length 1985;
Best Local Similarity 96.4%;   Pred. No. 3.4e-265;
  Matches 570; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

Qy       5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           |||||||||||||||||||||||||||||:|:|||||||| :||||||||||||||||||
Db    1395 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 1454

Qy      65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR 124
           |||||||||||||||||||||||||||||:||||||||||||||||||:||||||||||:|
Db    1455 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIH 1514

Qy     125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1515 SVWKDLLEDTVPIDTTIMAKQNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 1574

Qy     185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           |||||||||||||||||:|||||||||:||||::||||||||||||||||||||||||||
Db    1575 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 1634

Qy     245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db    1635 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 1694

Qy     305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1695 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 1754

Qy     365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1755 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 1814

Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1815 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 1874

Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db    1875 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 1934

Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||:||||||||||||||||||||||||||||||||||||||:|||||||||
Db    1935 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 1985
```

RESULT 7
US-10-471-164-5
; Sequence 5, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471,164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 591
; TYPE: PRT
; ORGANISM: HCV
US-10-471-164-5

```
Query Match          95.1%;   Score 3011; DB 16;  Length 591;
Best Local Similarity 95.9%;   Pred. No. 7.1e-266;
  Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

Qy       5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           |||||||||||||||||||||||||||||:|:|||||| :||||||||||||||||||||
Db       1 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 60
```

(Left rotated column)

```
Db       3 SMSYTWTGALITPCAAEEQKLPINALSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 62
           |||||||||||||||||:|||||||||||||||||:|||:||||||||||||||||||||
Qy      65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR 124
           |||::|||||||||||||||||||||||||||||||||||||||||||||:||:||||||
Db      63 DDHYQDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIH 122

Qy     125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 SVWKDLLEDTPIDTTIMAKQNEVFCVQPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTV 182

Qy     185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           |||||||||||||||||:|||||||||:|||::|||||||||||||||||||||||||||
Db     183 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRVEESIY 242

Qy     245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db     243 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 302

Qy     305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 362

Qy     365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db     363 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 422

Qy     425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 482

Qy     485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||||||||||||||||||||||:||||||||||||:|||||||||
Db     483 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 542

Qy     545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
           |||:||||||||||||||||||||||||||||||||||||||||||||||
Db     543 AASQLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 593
```

RESULT 6
US-10-492-178-6
; Sequence 6, Application US/10492178
; Publication No. US20040247615A1
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Kaslow, David C.
; APPLICANT: Bett, Andrew J.
; APPLICANT: Shiver, John W.
; APPLICANT: Nicosia, Alfredo
; APPLICANT: Lahm, Armin
; APPLICANT: Luzzago, Alessandra
; APPLICANT: Cortese, Riccardo
; APPLICANT: Colloca, Stefano
; TITLE OF INVENTION: HEPATITIS C VIRUS VACCINE
; FILE REFERENCE: ITR0015YP
; CURRENT APPLICATION NUMBER: US/10/492,178
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: PCT/US02/32512
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/363,774
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/328,655
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: NS sequence
US-10-492-178-6

```
Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
         ||||||||||||||||||||||||||:||||||||||||||||||||||:||::|:||
Db    61 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 120

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
         |||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db   121 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFIVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
         ||||||||||||||||||:||||:||:||||:||||||||||||||||||:|||||||||
Db   181 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 240

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
         |||||||||||||:|||||||:|||:||||||||||||||||||||||||||||||||:
Db   241 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
         |||||||||||||||||||||||||||||||||:|||||||||||||||:|:||||||
Db   301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 360

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
         ||||||||||||||||||||||||||||:||||||||||||||:||:||||||||||||
Db   361 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 420

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
         |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
         ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   481 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 540

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
         ||||||||||||||:|||||||||||||||||||||||||||||||||||
Db   541 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 591
```

```
Query Match   95.1%;  Score 3011;  DB 9;  Length 621;
Best Local Similarity   95.9%;  Pred. No. 7.6e-266;
Matches   567;  Conservative   18;  Mismatches   6;  Indels   0;  Gaps   0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
         ||||||||||||||||||:|||||||||||:|||:|:|||||||||:||||||||||||||
Db    31 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 90

Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
         ||||||||||||||||||||||||||:||||||||||||||||||||||:||::|:||
Db    91 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 150
```

```
Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
         |||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db   151 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 210

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFIVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
         ||||||||||||||||||:||||:||:||||:||||||||||||||||||:|||||||||
Db   211 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
         |||||||||||||:|||||||:|||:||||||||||||||||||||||||||||||||:
Db   271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 330

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
         |||||||||||||||||||||||||||||||||:|||||||||||||||:|:||||||
Db   331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 390

Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
         ||||||||||||||||||||||||||||:||||||||||||||:||:||||||||||||
Db   391 ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 450

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
         |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   451 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 510

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
         ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   511 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 570

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
         ||||||||||||||:|||||||||||||||||||||||||||||||||||
Db   571 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 621
```

```
Query Match   95.1%;  Score 3011;  DB 14;  Length 621;
Best Local Similarity   95.9%;  Pred. 7.6e-266;
Matches   567;  Conservative   18;  Mismatches   6;  Indels   0;  Gaps   0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
         ||||||||||||||||||:|||||||||||:|||:|:|||||||||:||||||||||||||
Db    31 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 90

Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEERACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
         ||||||||||||||||||||||||||:||||||||||||||||||||||:||::|:||
Db    91 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 150

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
         |||||||||:|||:||||||||||||||||||||||||||||||||||||||||||||
Db   151 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 210

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
         ||||||||||||||||||:||||||||:||||:||||||||||||||||||:|||||||||
Db   185 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 244
```

```
Db  211 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy  245 QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
        ||||||||||||:||||||||:||||:||||||||||||||||||||||||||||||||
Db  271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 330
Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
        ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
Db  331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 390
Qy  365 ELITSCSSNVSVAHDASGKRVVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
        |||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||
Db  391 ELITSCSSNVSVAHDASGKRVVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 450
Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
        |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  451 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 510
Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
        |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db  511 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 570
Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
        ||||||||||||||:|||||||||||||||||||||||||||||||||||
Db  571 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 621
```

RESULT 10
US-10-238-282-1
; Sequence 1, Application US/10238282
; Publication No. US20030232816A1
; GENERAL INFORMATION:
;   APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
;   TITLE OF INVENTION: VIRAL POLYMERASE INHIBITORS
;   FILE REFERENCE: 13/079
;   CURRENT APPLICATION NUMBER: US/10/238,282
;   CURRENT FILING DATE: 2002-09-10
;   PRIOR APPLICATION NUMBER: 60/216,084
;   PRIOR FILING DATE: 2000-07-06
;   PRIOR APPLICATION NUMBER: 60/274,374
;   PRIOR FILING DATE: 2001-03-08
;   PRIOR APPLICATION NUMBER: 60/281,434
;   PRIOR FILING DATE: 2001-04-02
;   NUMBER OF SEQ ID NOS: 1
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 621
;   TYPE: PRT
;   ORGANISM: Hepatitis C Virus
US-10-238-282-1

```
Query Match             95.1%;    Score 3011;   DB 15;   Length 621;
Best Local Similarity   95.9%;    Pred. No. 7.6e-266;
Matches  567;  Conservative  18;  Mismatches  6;  Indels  0;  Gaps  0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
        |||||||||||||||||||:||||||||||:|||:|:||||||||:|||||||||||||||
Db   31 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 90
Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
        ||||||||||||||||||||||||||||||||||||||||||||||||:||:||:|||||
Db   91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 150
Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
        ||||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||
Db  151 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 210
Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
        ||||||||||||||||||||||||||:||:||||||||||||||||||||:|||||||||
Db  211 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy  245 QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
        ||||||||||||:||||||||:||||:|||||||||||||||||||||||||||||||||
```

```
Db  271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 330
Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
        ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
Db  331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 390
Qy  365 ELITSCSSNVSVAHDASGKRVVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
        |||||||||||||||||||||||||||||:|||||||||||||:||||||||||||||||
Db  391 ELITSCSSNVSVAHDASGKRVVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 450
Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
        |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  451 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 510
Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
        |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db  511 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 570
Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
        ||||||||||||||:|||||||||||||||||||||||||||||||||||
Db  571 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 621
```

RESULT 11
US-10-198-259A-1
; Sequence 1, Application US/10198259A
; Publication No. US2003023625lA1
; GENERAL INFORMATION:
;   APPLICANT: Boehringer Ingelheim (Canada) Ltd.
;   TITLE OF INVENTION: Viral Polymerase Inhibitors
;   FILE REFERENCE: 13/089
;   CURRENT APPLICATION NUMBER: US/10/198,259A
;   CURRENT FILING DATE: 2002-07-18
;   PRIOR APPLICATION NUMBER: 60/306,669
;   PRIOR FILING DATE: 2001-07-20
;   PRIOR APPLICATION NUMBER: 60/338,324
;   PRIOR FILING DATE: 2001-12-07
;   NUMBER OF SEQ ID NOS: 4
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 621
;   TYPE: PRT
;   ORGANISM: HCV NS5B
US-10-198-259A-1

```
Query Match             95.1%;    Score 3011;   DB 15;   Length 621;
Best Local Similarity   95.9%;    Pred. No. 7.6e-266;
Matches  567;  Conservative  18;  Mismatches  6;  Indels  0;  Gaps  0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
        |||||||||||||||||||:||||||||||:|||:|:||||||||:|||||||||||||||
Db   31 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 90
Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
        ||||||||||||||||||||||||||||||||||||||||||||||||:||:||:|||||
Db   91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 150
Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
        ||||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||
Db  151 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 210
Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
        ||||||||||||||||||||||||||:||:||||||||||||||||||||:|||||||||
Db  211 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270
Qy  245 QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
        ||||||||||||:||||||||:||||:|||||||||||||||||||||||||||||||||
Db  271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 330
Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
        ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
Db  331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 390
```

```
Qy  365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||||||||||||||||||:|||||||||||:|||||||||||||||
Db  391  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW  450

Qy  425  ARMILMTHFFSILLAQBQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  451  ARMVLMTHFFSILLAQBQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  510

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db  511  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  570

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR  595
         |||||||||||||||:||||||||||||||||||||||||||||||||||
Db  571  AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR  621
```

RESULT 12
US-10-198-384A-1
; Sequence 1, Application US/10198384A
; Publication No. US2004024190A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: Viral Polymerase inhibitors
; FILE REFERENCE: 13/090
; CURRENT APPLICATION NUMBER: US/10/198,384A
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/338,061
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/307,674
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 621
;   TYPE: PRT
;   ORGANISM: HCV NS5B
US-10-198-384A-1

Query Match   95.1%; Score 3011; DB 15; Length 621;
Best Local Similarity 95.9%; Pred. No. 7.6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         |||||||||||||||||||:|||||||||||:||:|:|||||||||::||||||||||||
Db   31  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL  90

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         ||||||||||||||||||||||||||||||||||||||||||||||||||:||:||:|||
Db   91  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  150

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         ||||||||||:||:|||||||||||||||||||||||||||||||||||||||||||||
Db  151  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  210

Qy  185  TLPQAVMGSSYGFQYSPKQRVBFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         |||||||||||||||||||||||||||:|||:||||||||||||||||||:||||||||
Db  211  TLPQAVMGSSYGFQYSPKQRVBFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  270

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         ||||||||||||:||||||||:|||:||||||||||||||||||||||||||||||||:
Db  271  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  330

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
Db  331  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL  390

Qy  365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db  391  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW  450

Qy  425  ARMILMTHFFSILLAQBQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||:
```

```
Db  451  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  510

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  511  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  570

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR  595
         |||||||||||||||:||||||||||||||||||||||||||||||||||
Db  571  AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR  621
```

RESULT 13
US-10-471-164-1
; Sequence 1, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY.
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471,164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 621
;   TYPE: PRT
;   ORGANISM: HCV
US-10-471-164-1

Query Match   95.1%; Score 3011; DB 16; Length 621;
Best Local Similarity 95.9%; Pred. No. 7.6e-266;
Matches 567; Conservative 18; Mismatches 6; Indels 0; Gaps 0;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         |||||||||||||||||||:|||||||||||:||:|:|||||||||::||||||||||||
Db   31  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL  90

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         ||||||||||||||||||||||||||||||||||||||||||||||||||:||:||:|||
Db   91  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  150

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         ||||||||||:||:|||||||||||||||||||||||||||||||||||||||||||||
Db  151  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  210

Qy  185  TLPQAVMGSSYGFQYSPKQRVBFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         |||||||||||||||||||||||||||:|||:||||||||||||||||||:||||||||
Db  211  TLPQAVMGSSYGFQYSPKQRVBFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  270

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         ||||||||||||:||||||||:|||:||||||||||||||||||||||||||||||||:
Db  271  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  330

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||||||||||||||||||||||:||||||||||||||||||:||||||
Db  331  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL  390

Qy  365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||||||||||||||:|||||||||||:||||||||||||||||||
Db  391  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW  450

Qy  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  451  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  510

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  511  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  570

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLSVGVGIYLLPNR  595
```

Db   571 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 621

RESULT 14
US-10-851-710-1
; Sequence 1, Application US/10851710
; Publication No. US20040024955A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: VIRAL POLYMERASE INHIBITORS
; FILE REFERENCE: 13/079
; CURRENT APPLICATION NUMBER: US/10/851,710
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/10/238,282
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/216,084
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/281,434
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-851-710-1

Query Match   95.1%; Score 3011; DB 16; Length 621;
Best Local Similarity   95.9%; Pred. No. 7.6e-266;
Matches   567; Conservative   18; Mismatches   6; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||| :|::||||||||||||| : ||||||||||||||||
Db   31 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 90

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          ||||||||||||||||||||||||||||||||||||||||||||:|||||:|:||||||||
Db   91 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 150

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   151 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 210

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||||||||:|||| :|||||:|||||||||||||||||||||||||
Db   211 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 270

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   271 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 330

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||||| :|||||||||||||||| ||||||
Db   331 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 390

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   391 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 450

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   451 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 510

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   511 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 570

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   571 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 621

RESULT 15
US-09-838-386-12
; Sequence 12, Application US/09838386
; Patent No. US20010055756A1
; GENERAL INFORMATION:
; APPLICANT: Pellerin, Charles
; APPLICANT: Kukolj, George
; TITLE OF INVENTION: Internal De No. US20010055756Alo Initiation Sites of the HCV NS5B
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 1011.2180001
; CURRENT APPLICATION NUMBER: US/09/838,386
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,793
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: HTaA5B polymerase
US-09-838-386-12

Query Match   95.1%; Score 3011; DB 9; Length 627;
Best Local Similarity   95.9%; Pred. No. 7.7e-266;
Matches   567; Conservative   18; Mismatches   6; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||| :|::||||||||||||| : ||||||||||||||||
Db   37 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 96

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          ||||||||||||||||||||||||||||||||||||||||||||:|||||:|:||||||||
Db   97 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 156

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   157 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 216

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          |||||||||||||||||||||||:|||| :|||||:|||||||||||||||||||||||||
Db   217 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 276

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   277 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 336

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          ||||||||||||||||||||||||||||||||||| :|||||||||||||||| ||||||
Db   337 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 396

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   397 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW 456

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   457 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 516

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   517 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 576

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 595
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   577 AASRLDLSGWFVAGYNGGDIYHSLSRARPRWFMLCLLLLSVGVGIYLLPNR 627

Search completed: September 22, 2005, 14:49:31
Job time : 79 secs

us-10-712-479-2.rapb

This Page Blank (uspto)

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        September 22, 2005, 14:44:29 ; Search time 78 Seconds
                                    (without alignments)
                               2875.911 Million cell updates/sec

Title:          US-10-712-479-4
Perfect score:  3055
Sequence:       1 MASMSMSYTWTGALITPCAA.........GGDIYHSLSRARPRHHHHHH 580

Scoring table:  BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:       2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:          1964263

Minimum DB seq length: 0
Maximum DB seq length: 585

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_16Dec04:*
                1: geneseqp1980s:*
                2: geneseqp1990s:*
                3: geneseqp2000s:*
                4: geneseqp2001s:*
                5: geneseqp2002s:*
                6: geneseqp2003as:*
                7: geneseqp2003bs:*
                8: geneseqp2004s:*

            Pred. No. is the number of results predicted by chance to have a
            score greater than or equal to the score of the result being printed,
            and is derived by analysis of the total score distribution.

                           SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3055	100.0	580	8	ADO19023	ADO19023 C-termina
2	2951	96.6	576	6	AAO26417	AAO26417 Mutant RN
3	2948	96.5	576	6	AAO26419	AAO26419 Mutant RN
4	2944	96.4	576	6	AAO26418	AAO26418 Mutant RN
5	2942	96.3	579	4	AAB60131	AAB60131 Hepatitis
6	2936	96.1	578	5	AAG79558	AAG79558 HCV decre
7	2936	96.1	578	6	ABP71662	ABP71662 HCV NS5Bd
8	2879	94.2	576	8	ABG72214	ABG72214 HCV mutan
9	2879	94.2	576	8	ADH74494	ADH74494 HCV RNA-d
10	2846	93.2	567	7	ADM69007	ADM69007 HCV NS5B
11	2734.5	89.5	545	2	AAG79559	AAG79559 HCV decre
12	2279	74.6	462	2	AAR58592	AAR58592 Hepatitis
13	2279	74.6	462	8	ADL17783	ADL17783 Hepatitis
14	2144	70.2	547	2	AAR67630	AAR67630 Non-A Non
15	1812	59.3	389	2	AAR29906	AAR29906 HCV NS4-N
16	1801	59.0	389	2	AAR29904	AAR29904 HCV NS4-N
17	1798	58.9	389	2	AAR29905	AAR29905 HCV NS4-N
18	1798	58.9	389	2	AAR29876	AAR29876 HCV NS4-N
19	1755	57.4	365	2	AAR38284	AAR38284 NANB hepa
20	1645	53.8	365	2	AAR38283	AAR38283 NANB hepa
21	1634	53.5	365	2	AAR67592	AAR67592 Hepatitis
22	1630	53.4	365	2	AAR67590	AAR67590 Hepatitis
23	1430	46.8	365	2	AAR50081	AAR50081 NANBH vir
24	1430	46.8	352	8	ADJ81686	ADJ81686 Non-A-non
25	1391	45.5	300	5	ABB77283	ABB77283 HCV bait

26	1391	45.5	365	2	AAR38287	AAR38287 NANB hepa
27	1384	45.3	365	2	AAR38286	AAR38286 NANB hepa
28	1375	45.0	365	2	AAR38285	AAR38285 NANB hepa
29	1363	44.6	285	2	AAR29879	AAR29879 HCV NS4-N
30	1232	40.3	260	5	ABB77285	ABB77285 HCV bait
31	1178	38.6	270	2	AAR25896	AAR25896 HK21. 9/2
32	1177	38.5	247	2	AAR25874	AAR25874 HCV polyp
33	1177	38.5	247	2	AAW41752	AAW41752 Hepatitis
34	1175	38.5	239	2	AAR25869	AAR25869 HCV polyp
35	1175	38.5	239	2	AAW41747	AAW41747 Hepatitis
36	1175	38.5	253	2	AAR25891	AAR25891 HK16. 9/2
37	1158.5	37.9	251	2	AAR25895	AAR25895 HK20. 9/2
38	1154	37.8	224	2	AAR25873	AAR25873 HCV polyp
39	1154	37.8	224	2	AAW41751	AAW41751 Hepatitis
40	1129	37.0	496	2	AAR21568	AAR21568 HCV CKS-N
41	1124	36.8	496	2	AAR33635	AAR33635 HCV CKS-N
42	1124	36.8	496	2	AAR33597	AAR33597 HCV CKS-N
43	1124	36.8	496	2	AAR33577	AAR33577 HCV CKS-N
44	1124	36.8	496	4	AAB51375	AAB51375 HCV recom
45	1124	36.8	496	7	ABW01861	ABW01861 HCV-CKS-N

```
                           ALIGNMENTS

RESULT 1
ADO19023
ID   ADO19023 standard; protein; 580 AA.
XX
AC   ADO19023;
XX
DT   12-AUG-2004  (first entry)
XX
DE   C-terminally truncated HCV NS5B polymerase, Cdelta21 NS5B.
XX
KW   RNA polymerase activity; continuous-read assay;
KW   Hepatitis C virus infection; HCV infection; polymerase; NS5B;
KW   bone mineral disease; osteoporosis; carcinoma; cardiovascular disease;
KW   diabetes; ocular disorder; renal dysfunction; lymphoma;
KW   lymphoproliferative disorder; metabolic disorder; arthritis;
KW   sleep disorder; thyroid disorder; antiinflammatory; hepatotropic;
KW   virucide; osteopathic; cytostatic; antidiabetic; ophthalmological;
KW   nephrotropic; antiarthritic; enzyme.
XX
OS   Hepatitis C virus.
XX
PN   WO200404228-A2.
XX
PD   27-MAY-2004.
XX
PF   13-NOV-2003; 2003WO-US036465.
XX
PR   13-NOV-2002; 2002US-0425981P.
XX
PA   (PFIZ ) PFIZER INC.
XX
PI   Yagi Y, Sheets MP, Wells PA, Shelly JA, Poorman RA, Epps DE;
PI   Morgan AG;
XX
DR   WPI; 2004-420337/39.
DR   N-PSDB; ADO19022.
XX
PT   Detecting RNA polymerase activity in a continuous-read manner, useful in
PT   treating osteoporosis, carcinomas, cardiovascular diseases, ocular
PT   disorders or arthritis, by contacting an RNA polymerase with an
PT   oligonucleotide template.
XX
PS   Claim 7; SEQ ID NO 4; 46pp; English.
XX
CC   The present invention relates to a method for detecting RNA polymerase
CC   activity in a continuous-read manner. The method comprises contacting an
CC   RNA polymerase with an oligonucleotide template in a reaction mixture
CC   comprising an assay buffer, under conditions in which the RNA polymerase
```

```
CC   is active, adding a fluorescent dye capable of binding double-stranded
CC   nucleic acid molecules to the reaction mixture, and measuring the
CC   fluorescence of the reaction mixture. The RNA polymerase is the Hepatitis
CC   C virus (HCV) polymerase, NS5B. The method is useful for treating HCV
CC   infection, bone mineral diseases like osteoporosis, carcinomas,
CC   cardiovascular diseases, diabetes, ocular disorders, renal dysfunction,
CC   lymphomas, lymphoproliferative disorders, metabolic disorders, arthritis,
CC   sleep disorders and thyroid disorders. The present sequence represents C-
CC   terminally truncated HCV NS5B polymerase (designated Cdelta21 NS5B).
XX
SQ   Sequence 580 AA;

Query Match             100.0%;  Score 3055;  DB 8;  Length 580;
Best Local Similarity   100.0%;  Pred. No. 7e-285;
Matches   580;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1  MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDR   60

Qy   61  LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV  120

Qy  121  NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY  180

Qy  181  DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVE  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVE  240

Qy  241  ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY  300

Qy  301  LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP  360

Qy  361  EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYA  420

Qy  421  PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS  480

Qy  481  YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL  540

Qy  541  TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  580
         ||||||||||||||||||||||||||||||||||||||||
Db  541  TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  580

RESULT 2
AAO26417
ID   AAO26417 standard; protein; 576 AA.
XX
AC   AAO26417;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Mutant RNA polymerase MUT-2 protein.
XX
KW   Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.
XX
OS   Unidentified.
XX
PN   EP1256628-A2.
XX
```

```
PD   13-NOV-2002.
XX
PF   07-MAY-2002; 2002EP-00009387.
XX
PR   10-MAY-2001; 2001US-0289829P.
XX
PA   (AGOU-) AGOURON PHARM INC.
XX
PI   Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;
XX
DR   WPI; 2003-031968/03.
XX
PT   Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA
PT   polymerase, useful for producing the mutant RNA polymerase which can be
PT   used for identifying inhibitors that can treat HCV infection.
XX
PS   Claim 7; Page; 32pp; English.
XX
CC   The invention relates to an isolated, purified nucleic acid molecule,
CC   which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where
CC   a point mutation occurs at least one or all of positions 47, 101, 106,
CC   and/or 114 of the native RNA polymerase sequence not defined in the
CC   specification. The method of the invention is utilised to identify
CC   inhibitors of the polymerase. The inhibitors serving as lead compounds
CC   for the design of potentially therapeutic compounds for the treatment of
CC   HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation
CC   properties as compared to a native HCV NS5B RNA polymerase. This sequence
CC   represents a mutant RNA polymerase protein of the invention. NOTE: This
CC   sequence is not shown in the specification. It is, however, described in
CC   the specification and its sequence has been obtained from an electronic
CC   data file
XX
SQ   Sequence 576 AA;

Query Match             96.6%;  Score 2951;  DB 6;  Length 576;
Best Local Similarity   96.5%;  Pred. No. 7.4e-275;
Matches   556;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
         |||||||||||||||||||||||||||||||||||:||| |||||||:||||||||||||
Db    1  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL   60

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||:
Db   61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIH  120

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         ||||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||
Db  121  SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         ||||:||||||||||||:|||||||||||:||:|||||||||||||||||||||||||||
Db  181  TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY  240

Qy  245  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||:
Db  241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  360

Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  361  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  420

Qy  425  ARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         |||||||||||||||||||||||||||||:||||||||||:||||||||||:||||||||
Db  481  EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP  540
```

us-10-712-479-4.szlm585.rag

```
Qy   181 TLPQVVMGSSYGFQYSPSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY  240
         |||||||||| ||| :| |||||||        ||||||||||||||||||||||||||||||
Db   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA    304

Qy   241 QCCDLAPEARQAIKSLTERLYIIGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS    300
         |||||||||||| |||||||| |||||||||||||||||||||||||||||||||||||:
Db   301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  360

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  420

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP  540

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||| |||||||||| |||||||| ||||||||||
Db   541 AASQLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          576

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          580
         ||:|||||||||||||||||||||||||||||||
Db   541 AASQLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          576


RESULT 4
AAO26418
ID   AAO26418 standard; protein; 576 AA.
XX
AC   AAO26418;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Mutant RNA polymerase MUT-1 protein.
XX
KW   Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.
XX
OS   Unidentified.
XX
PN   EP1256628-A2.
XX
PD   13-NOV-2002.
XX
PF   07-MAY-2002; 2002EP-00009387.
XX
PR   10-MAY-2001; 2001US-0289829P.
XX
PA   (AGOU-) AGOURON PHARM INC.
XX
PI   Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;
XX
DR   WPI; 2003-031968/03.
XX
PT   Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA
PT   polymerase, useful for producing the mutant RNA polymerase which can be
PT   used for identifying inhibitors that can treat HCV infection.
XX
PS   Claim 8; Page; 32pp; English.
XX
CC   The invention relates to an isolated, purified nucleic acid molecule,
CC   which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where
CC   a point mutation occurs at least at one or all of positions 47, 101, 106,
CC   and/or 114 of the native RNA polymerase sequence not defined in the
CC   specification. The method of the invention is utilised to identify
CC   inhibitors of the polymerase. The inhibitors serving as lead compounds
CC   for the design of potentially therapeutic compounds for the treatment of
CC   HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation
CC   properties as compared to a native HCV NS5B RNA polymerase. This sequence
CC   represents a mutant RNA polymerase protein of the invention. NOTE: This
CC   sequence is not shown in the specification. It is, however, described in
```

```
Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          580
         ||:|||||||||||||||||||||||||||||||
Db   541 AASQLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          576


RESULT 3
AAO26419
ID   AAO26419 standard; protein; 576 AA.
XX
AC   AAO26419;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Mutant RNA polymerase MUT-5 protein.
XX
KW   Antiviral; hepatitis C virus; HCV; NS5B RNA polymerase.
XX
OS   Unidentified.
XX
PN   EP1256628-A2.
XX
PD   13-NOV-2002.
XX
PF   07-MAY-2002; 2002EP-00009387.
XX
PR   10-MAY-2001; 2001US-0289829P.
XX
PA   (AGOU-) AGOURON PHARM INC.
XX
PI   Love RA, Yu X, Diehl W, Hickey MJ, Parge HE, Gao J, Fuhrman S;
XX
DR   WPI; 2003-031968/03.
XX
PT   Isolated nucleic acid encoding a mutant hepatitis C virus (HCV) NS5B RNA
PT   polymerase, useful for producing the mutant RNA polymerase which can be
PT   used for identifying inhibitors that can treat HCV infection.
XX
PS   Claim 9; Page; 32pp; English.
XX
CC   The invention relates to an isolated, purified nucleic acid molecule,
CC   which encodes a mutant hepatitis C virus (HCV) NS5B RNA polymerase, where
CC   a point mutation occurs at least at one or all of positions 47, 101, 106,
CC   and/or 114 of the native RNA polymerase sequence not defined in the
CC   specification. The method of the invention is utilised to identify
CC   inhibitors of the polymerase. The inhibitors serving as lead compounds
CC   for the design of potentially therapeutic compounds for the treatment of
CC   HCV. The mutant HCV NS5B RNA polymerase has improved crystallisation
CC   properties as compared to a native HCV NS5B RNA polymerase. This
CC   sequence is not shown in the specification. It is, however, described in
CC   the specification and its sequence has been obtained from an electronic
CC   data file
XX
SQ   Sequence 576 AA;

Query Match             96.5%;  Score 2948;  DB 6;  Length 576;
Best Local Similarity   96.4%;
Matches  555;  Conservative  12;  Mismatches  9;  Indels  0;  Gaps  0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
         ||||||||||||||||||||||||||||||||||||| |:|| |||||||||||||||||
Db     1 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL   60

Qy    65 DDHYRDVLKEMCKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR  124
         ||||||||||||||||||||||||||||||||||||||||||||| |:|| |||:|||||||
Db    61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKYGYGAKDVRNLSSRAVNHIH  120

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         |||||||||||| ||:|||||||||||||||||||||||||||||||||||||||||||||
Db   121 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
```

CC   the specification and its sequence has been obtained from an electronic
CC   data file
XX
SQ   Sequence 576 AA;

Query Match   96.4%; Score 2944; DB 6; Length 576;
Best Local Similarity   96.2%; Pred. No. 3.5e-274;
Matches 554; Conservative   13; Mismatches   9; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
              |||||||||||||||||||||||||||||||||:||:|:|||||||||||||||||||
Db   1 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVL 60

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
              |||||||||||||||||||||||||||||||||||||||||||||:|||:|||:||||||
Db   61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQDVRNLSSKAVNHIH 120

Qy   125 SVWKDLLEDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
              ||||||||||:|:||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
              ||||:||||||||||||:|||||||||||:||:|||||||||||||||||||||||||||
Db   181 TLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIY 240

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
              ||||||||||||:|||||||:|||:|||||||||||||||||||||||||||||||||||
Db   241 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 360

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 420

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
              ||||||||||||||||||||||||||||||:||||||||:|||||||||:||||||||||
Db   481 EINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIP 540

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH 580
              |||:||||||||||||||||||||||||||||||
Db   541 AASQLDLSGWFVAGYSGGDIYHSLSRARPRGHHHHH 576

RESULT 5
AAB60131
ID   AAB60131 standard; protein; 579 AA.
XX
AC   AAB60131;
XX
DT   03-APR-2001 (first entry)
XX
DE   Hepatitis C virus NS5B polymerase related protein.
XX
KW   Hepatitis C virus; HCV; NS5B polymerase; protein coordinate data;
KW   cirrhosis; hepatoma; inhibitor.
XX
OS   Hepatitis C virus.
XX
PN   EP1065213-A2.
XX
PD   03-JAN-2001.
XX
PF   30-JUN-2000; 2000EP-00113955.
XX
PR   02-JUL-1999; 99JP-00188630.
PR   07-JUL-1999; 99JP-00192488.

XX
PA   (NISB) JAPAN TOBACCO INC.
XX
PI   Ago H, Miyano M, Adachi T;
XX
DR   WPI; 2001-104834/12.
XX
PT   New polypeptide, derived from hepatitis C virus (HCV) polymerase NS5B,
PT   having a HCV polymerase activity, useful for crystal structure analysis
PT   and for rational identification of HCV polymerase inhibitors.
XX
PS   Disclosure; Page 279-282; 295pp; English.
XX
CC   The present invention provides the hepatitis C virus (HCV) NS5B
CC   polymerase protein and its structure. It also provides methods for
CC   identifying inhibitors of the protein which can be used in the treatment
CC   of HCV infection. HCV can lead to cirrhosis and hepatoma in its chronic
CC   form
XX
SQ   Sequence 579 AA;

Query Match   96.3%; Score 2942; DB 4; Length 579;
Best Local Similarity   96.0%; Pred. No. 5.5e-274;
Matches 556; Conservative   12; Mismatches   12; Indels   9; Gaps   1;

Qy   4 MSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQV 63
              ||||||||||||||||||||||||||||||||||||:||:|:|||||||||||||||||
Db   1 MSMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQV 60

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHI 123
              |||||||||||||||||||||||||||||||||||||||||||||:|||:|||:|||||
Db   61 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAQDVRNLSSKAVNHI 120

Qy   124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 183
              :|||||||||::||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 HSVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 180

Qy   184 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
              |||||:||||||||||||:|||||||||||:||:||||||||||||||||||||||||||
Db   181 STLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESI 240

Qy   244 YQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 303
              |||||||||||||:|||||||:|||:||||||||||||||||||||||||||||||||||
Db   241 YQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 300

Qy   304 AAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYD 363
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 SAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYD 360

Qy   364 LELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTL 423
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   361 LELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTL 420

Qy   424 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 WARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSP 480

Qy   484 GEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPI 543
              ||||||||||||||||||||||||||||||:||||||||:|||||||||:||||||||||
Db   481 GEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPI 540

Qy   544 PAASRLDLSGWFVAGYSGGDIYHSLSRARPR--HHHHHH 580
              ||||:|||||||||||||||||||||||||||:||||||
Db   541 PAASQLDLSGWFVAGYSGGDIYHSLSRARPRGHHHHH 579

RESULT 6
AAG79558
ID   AAG79558 standard; protein; 578 AA.
XX
AC   AAG79558;
XX
DT   09-DEC-2002 (first entry)

```
XX
DE   HCV decrease-affinity NS5B polymerase, NS5Bdelta21C-HT.
XX
KW   Enzyme; hepatitis C virus; HCV; decreased affinity; NS5B; polymerase;
KW   inhibitor; RNA-dependent RNA polymerase.
XX
OS   Hepatitis C virus.
XX
PN   WO20020739-A2.
XX
PD   12-SEP-2002.
XX
PF   06-MAR-2002; 2002WO-CA000323.
XX
PR   08-MAR-2001; 2001US-0274374P.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX
PI   Kukolj G, Mckercher G;
XX
DR   WPI; 2002-698759/75.
XX
PT   Identifying inhibitors HCV NS5B RNA-dependent RNA polymerase, comprises
PT   incubating an HCV NS5B polymerase having decreased affinity for the
PT   primer template relative to the native polymerase, with a potential
PT   inhibitor.
XX
XX   Claim 9; Page 47-48; 52pp; English.
XX
CC   The sequences given in AAG79556-61 are hepatitis C virus (HCV) decreased
CC   affinity NS5B polymerases. These enzymes are used in the method of the
CC   invention for identifying a potential inhibitor of the binding between a
CC   HCV NS5B RNA-dependent RNA polymerase and an appropriate primer-template.
CC   The method comprises incubating the HCV NS5B polymerase with the primer-
CC   template in the presence and absence of a potential inhibitor. The HCV
CC   NS5B polymerase has a decreased affinity for the primer-template relative
CC   to that of native HCV NS5B RNA-dependent RNA polymerase. The method is
CC   useful for identifying a potential inhibitor of the binding between a HCV
CC   NS5B RNA-dependent RNA polymerase and a primer-template. Use of
CC   polymerase constructs having a lower affinity towards the primer-template
CC   than that of native NS5B polymerase is particularly useful for
CC   identifying potential inhibitors in screening large libraries of
CC   compounds. The new method reduces the difficulties and disadvantages of
CC   prior art. The present method provides an assay that is easy to perform
CC   on large libraries of compounds, and has improved sensitivity for
CC   detecting inhibitors that would not be identified as such using native
CC   NS5B polymerase. This protein represents a soluble form of mature HCV
CC   NS5B which lacks the C-terminal 21 amino acids and has an C-terminal hexa
CC   -histidine tag
XX
SQ   Sequence 578 AA;
```

Query Match 96.1%; Score 2936; DB 5; Length 578;
Best Local Similarity 95.5%; Pred. No. 2.1e-273;
Matches 552; Conservative 18; Mismatches 6; Indels 2; Gaps 1;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKVTFDRLQVL   64
         ||||||||||||||||||||||||||||||:||||:||||||||||||||||||||||||
Db    1  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKVTFDRLQVL   60

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         |||||||||||||||||||||||||||:||||||||||||||||:||:||:|||||||||
Db   61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  120

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         |||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         ||||||||||||||||||||||||||:|:|||||||||||||||||:||:|||||||||:
Db  181  TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  240

Qy  245  QCCDLAPEARQAIRSLTERLYIVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
```

```
Db  241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300
         |||||||||||||:||||||||| ||:|||||||||||||||||||||||||||||||||:
Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db  301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL  360

Qy  365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||||||||||||||||||||||:|||||||:||||||||||||||||
Db  361  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW  420

Qy  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db  481  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  540

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRARPR--HHHHHH  580
         |||||||||||||||:||||||||||||||  ||||||
Db  541  AASRLDLSGWFVAGYNGGDIYHSLSRARPRLEHHHHHH  578
```

```
RESULT 7
ABP71662
ID   ABP71662 standard; protein; 578 AA.
XX
AC   ABP71662;
XX
DT   29-MAY-2003 (first entry)
XX
DE   HCV NS5Bdel21-His protein sequence.
XX
KW   HCV; polymerase; virucide; hepatotropic; antiviral; enzyme; NS5B.
XX
OS   Hepatitis C virus.
XX
PN   WO200314377-A2.
XX
PD   20-FEB-2003.
XX
PF   05-AUG-2002; 2002WO-CA001214.
XX
PR   07-AUG-2001; 2001US-0310272P.
XX
PA   (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX
PI   Kukolj G, Beaulieu PL, Mckercher G;
XX
DR   WPI; 2003-300535/29.
XX
PT   Identifying compounds that bind to hepatitis C virus polymerase, useful
PT   as antiviral agents, from ability to displace bound probe from the
PT   polymerase.
XX
PS   Example 2; Page 117-119; 125pp; English.
XX
CC   The invention relates to a method for identifying compounds that bind to
CC   hepatitis C virus (HCV) polymerase. The method involves (a) treating HCV
CC   polymerase, or its analogue, with a probe that binds to the polymerase to
CC   form a complex but is displaceable by an inhibitor; (b) measuring a
CC   signal emitted from the probe in the complex to establish a baseline
CC   value; (c) incubating the complex with test compound; and (d) measuring
CC   the signal from the complex and comparing it with the baseline. Any
CC   modulation indicates the test compound binds to HCV polymerase. The
CC   method is used to identify (also to determine potency) of HCV polymerase
CC   inhibitors, potentially useful as antiviral agents. The present sequence
CC   represents a HCV NS5Bdel121-His polymerase
XX
SQ   Sequence 578 AA;
```

Query Match 96.1%; Score 2936; DB 6; Length 578;

```
FT                   /note= "Preferred amino acid at this position is
FT                   glutamine. Wild type residue at this position is
FT                   arginine"
XX
PN   US6434489-B1.
XX
PD   13-AUG-2002.
XX
PF   03-APR-2000; 2000US-00541990.
XX
PR   03-APR-2000; 2000US-00541990.
PA   (SCHE ) SCHERING CORP.
XX
PI   Lesburg CA,  Cable M,  Hong Z,  Mannarino AF,  Weber PC;
XX
DR   WPI; 2003-110131/10.
XX
PT   New crystalline Hepatitis C virus NS5B polypeptide composition useful for
PT   determining the atomic coordinates of the polypeptide and drug screening.
XX
PS   Claim 1; Col 75-80; 44pp; English.
XX
CC   The present invention relates to the purification, crystallisation, and
CC   structure of hepatitis C virus (HCV) NS5B RNA-dependent RNA polymerase.
CC   The composition containing the crystals of HCV NS5B polymerase
CC   polypeptide diffracts x-rays for determination of the atomic coordinates
CC   of the polypeptide to greater than 4.0 Angstroms resolution. The crystal
CC   is useful for determination of the atomic coordinates of the HCV NS5B
CC   polypeptide, and is useful in drug screening methods. The present
CC   sequence represents a mutant of the published sequence (Genbank Accession
CC   No. 130458) for the HCV NS5B polypeptide, His-NS5BdelCT21. This mutant
CC   contains 3 amino acid substitutions and a polyhistidine tag at the N-
CC   terminus. The C-terminal 21 amino acids present in the wild type are also
CC   deleted in His-NS5BdelCT21
XX
SQ   Sequence 576 AA;

Query Match             94.2%;  Score 2879;  DB 6;  Length 576;
Best Local Similarity   96.0%;  Pred. No. 6.5e-268;
Matches 545;  Conservative 11;  Mismatches 12;  Indels 0;  Gaps 0;

Qy    7  SYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVLDD  66
         |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db    9  SYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDD  68

Qy   67  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSV  126
         ||||||||||||||||||||||||||||||||||||||||||||||:||:|:||:||:||
Db   69  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIHSV  128

Qy  127  WKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  186
         |||||||:|||| ::||||||||||||||||||||||||||||||||||||||||||||||
Db  129  WKDLLEDTVTPIPDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  188

Qy  187  PQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQC  246
         ||:|||||||||||||:|||||||||:||||:|||||||||||||||||||||||||||||
Db  189  PQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIYQC  248

Qy  247  CDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAA  306
         |||||||||:||||||||:|||| ||||||||||||||||||||||||||||||||:|||
Db  249  CDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAA  308

Qy  307  CRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLEL  366
         |||||||||||||||||||||||||| ||||||||: ||||||||||||||||||||||||
Db  309  CRAAKLQDCTMLVNGDDLVVICESAGXQEDAASLRXPTEAMTRYSAPPGDPPQPEYDLEL  368

Qy  367  ITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  426
         ||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||||
Db  369  ITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  428

Qy  427  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI  486
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  429  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI  488
```

```
Best Local Similarity   95.5%;  Pred. No. 2.1e-273;
Matches 552;  Conservative 18;  Mismatches 6;  Indels 2;  Gaps 1;

Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         ||||||||||||||||||||:|||||||||:||| |:|:||||:|||||||||||||||||
Db    1  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL  60

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         ||||||||||||||||||||||||||||||||||||||||||||||:||:|:||:||||||
Db   61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  120

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         ||||||||||:|||| :||||||||||||||||||||||||||||||||||||||||||||
Db  121  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         ||||||||||||||||||:|||||||:|||:|||||||||||||||||||:|||||||||
Db  181  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  240

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         ||||||||||||:||||||||:|||| |||||||||||||||||||||||||||||||||:
Db  241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300

Qy  305  ACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||||||||||||||||||||||:|||||||||||||||:|||||||||
Db  301  ACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAANLRVFTEAMTRYSAPPGDLPQPEYDL  360

Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         ||||||||||||||||||||||||||||| :||||||||||||:|||||||||||||||||
Db  361  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW  420

Qy  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db  481  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  540

Qy  545  AASRLDLSGWFVAGYSGDIYHSLSRARPR--HHHHHH  580
         ||||||||||||||||:|||||||||||||  ||||||
Db  541  AASRLDLSGWFVAGYNGGDIYHSLSRARPRLEHHHHHH  578
```

```
RESULT 8
ABG72214
ID   ABG72214 standard; protein; 576 AA.
XX
AC   ABG72214;
XX
DT   17-FEB-2003  (first entry)
XX
DE   HCV mutant NS5B polymerase polypeptide, His-NS5BdelCT21.
XX
KW   Crystallisation; crystal structure; hepatitis C virus; HCV;
KW   NS5B polymerase; NS5B RNA-dependent RNA polymerase; x-ray diffraction;
KW   atomic coordinate determination; drug screening; mutant; mutein;
KW   protein co-ordinate data; enzyme; His-NS5BdelCT21.
XX
OS   Hepatitis C virus; BK isolate.
OS   Synthetic.
XX
FH   Key              Location/Qualifiers
FT   Misc-difference  335
FT                    /label= Thr, Val
FT                    /note= "Preferred amino acid at this position is valine.
FT                    Wild type residue at this position is threonine"
FT   Misc-difference  344
FT                    /label= Val, Ala
FT                    /note= "Preferred amino acid at this position is alanine.
FT                    Wild type residue at this position is valine"
FT   Misc-difference  550
FT                    /label= Arg, Gln
```

```
Qy  487 NRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAA 546
            ||||||||||||||||||||||||||||||||| ||||||||||||||:||||||||||||
Db  489 NRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAA 548

Qy  547 SRLDLSGWFVAGYSGGDIYHSLSRARPR 574
            | ||||||||||||||||||||||||||
Db  549 SXLDLSGWFVAGYSGGDIYHSLSRARPR 576
```

```
RESULT 9
ADH74494
ID   ADH74494 standard; protein; 576 AA.
XX
AC   ADH74494;
XX
DT   25-MAR-2004  (first entry)
XX
DE   HCV RNA-dependent RNA polymerase NS5B polypeptide.
XX
KW   HCV; RNA-dependent RNA polymerase; NS5B; cation exchange chromatography;
KW   polyethylene glycol; crystallisation; tNS5B; truncated NS5B;
KW   anti-HCV agent.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 335
FT                   /label= Thr, Val
FT   Misc-difference 344
FT                   /label= Val, Ala
FT   Misc-difference 550
FT                   /label= Arg, Gln
XX
PN   US2003171874-A1.
XX
PD   11-SEP-2003.
XX
PF   12-JUN-2002; 2002US-00170131.
PR   03-APR-2000; 2000US-00541990.
XX
PA   (LESB/) LESBURG C A.
PA   (CABL/) CABLE M.
PA   (HONG/) HONG Z.
PA   (MANN/) MANNARINO A F.
PA   (WEBE/) WEBER P C.
XX
PI   Lesburg CA, Cable M, Hong Z, Mannarino AF, Weber PC;
XX
DR   WPI; 2004-119171/12.
XX
PT   Crystalline composition comprising a hepatitis C virus (HCV) RNA
PT   dependent RNA polymerase NS5b polypeptide, useful for providing a target
PT   for discovery of anti-HCV agents.
XX
PS   Claim 3; SEQ ID NO 1; 44pp; English.
XX
CC   The invention relates to a crystalline composition comprising a hepatitis
CC   C virus (HCV) RNA-dependent RNA polymerase NS5B polypeptide. The
CC   invention also relates to a composition comprising an NS5B protein having
CC   at least one amino acid substitution and a method of preparing purified
CC   crystalline HCV NS5B comprising stabilising an NS5B protein preparation
CC   in a solution containing a protein stabilising agent, subjecting the NS5B
CC   protein preparation to cation exchange chromatography in a buffer
CC   containing glycine, and allowing crystals to form in a precipitant
CC   containing a protein stabilising agent, a salt and polyethylene glycol
CC   under conditions in which crystallisation occurs. The NS5B polypeptide is
CC   preferably tNS5B, a truncated NS5B polypeptide. The invention is used to
CC   provide a target for discovery of anti-HCV agents. This sequence
CC   represents the NS5B polypeptide of the invention.
XX
SQ   Sequence 576 AA;
```

```
Query Match            94.2%;  Score 2879;  DB 8;  Length 576;
Best Local Similarity  96.0%;  Pred. No. 6.5e-268;
Matches  545; Conservative   11; Mismatches   12; Indels    0; Gaps    0;
```

```
Qy    7 SYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTSRSASLRQKKVTFDRLQVLDD 66
           ||||||||||||||||||||||||||||||||||:||||||:||||||||||||||||||
Db    9 SYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDD 68

Qy   67 HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSV 126
           |||||||||||||||||||||||||||||||||||||||||||||::||::|||||:|||
Db   69 HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIHSV 128

Qy  127 WKDLLEDTDTPIQTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL 186
           |||||||:|:|:||||||||||||||||||||||||||||||||||||||||||||||||
Db  129 WKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL 188

Qy  187 PQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQC 246
           ||.|||||||||||||||||||||||:||:::||||||||||||||||||||||||||||
Db  189 PQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIYQC 248

Qy  247 CDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAA 306
           ||||||||||:|||||||:||||.|||||||:||||||||||||||||||||||||:|.|
Db  249 CDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAA 308

Qy  307 CRAAKLQDCTMLVNGDDLVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLEL 366
           |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db  309 CRAAKLQDCTMLVNGDDLVICESAGXQEDAASLRXFTEAMTRYSAPPGDPPQPEYDLEL 368

Qy  367 ITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWAR 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  369 ITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR 428

Qy  427 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  429 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI 488

Qy  487 NRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAA 546
           ||||||||||||||||||||||||||| |||||||||:||||||||||||:||||||||||
Db  489 NRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAA 548

Qy  547 SRLDLSGWFVAGYSGGDIYHSLSRARPR 574
           | ||||||||||||||||||||||||||
Db  549 SXLDLSGWFVAGYSGGDIYHSLSRARPR 576
```

```
RESULT 10
ADM69007
ID   ADM69007 standard; protein; 567 AA.
XX
AC   ADM69007;
XX
DT   03-JUN-2004  (first entry)
XX
DE   HCV NS5B RNA-dependent RNA polymerase (RdRp) protein.
XX
KW   HCV; NS5B; RNA-dependent RNA polymerase; RdRp; growth inhibitor; enzyme.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 464
FT                   /note= "Encoded by GAA"
XX
PN   KR2002070125-A.
XX
PD   05-SEP-2002.
XX
PF   26-FEB-2002; 2002KR-00010174.
XX
PR   27-FEB-2001; 2001KR-00009873.
XX
PA   (LGLI-) LG LIFE SCI LTD.
```

XX
PI   Hyun US, Kim JM, Lee MG, Noh GY;
XX
DR   WPI; 2003-253755/25.
DR   N-PSDB; ADM69006.
XX
PT   Recombinant hepatitis C virus ns5b protein, and preparation process and
PT   use thereof.
XX
PS   Claim 3; SEQ ID NO 8; 19pp; Korean.
XX
CC   The invention relates to a novel recombinant Hepatitis C virus (HCV) NS5B
CC   protein which is derived from a HCV patient's serum and has excellent RNA
CC   -dependent RNA polymerase (RdRp) activity. The polypeptide of the
CC   invention may be useful as a HCV growth inhibitor. The current sequence
CC   is that of the HCV NS5B RNA-dependent RNA polymerase (RdRp) protein of
CC   the invention.
XX
SQ   Sequence 567 AA;

Query Match   93.2%; Score 2846; DB 7; Length 567;
Best Local Similarity   94.7%; Pred. No. 9.6e-265;
Matches 537; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
              ||||||||||||||||||||||||||||||||||||||||||:||:||||||||||||||
Db   1 SMSYSWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSAGLRQKEVTFDRLQVL 60

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
              ||||||||||||||||||||||||||||||||||||||||||||||||:|:||||||||
Db   61 DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIR 120

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db   121 SVWKDLLEDTETPINTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
              ||||||||||||||||||||||||||||:|||||||||||||||||||||:|||||||||
Db   181 TLPQAVMGSSYGFQYSPGQRVEFLVKAWKSKKCCPMGFSYDTRCFDSTVTESDIRVEESIY 240

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
              ||||||||:||||:|||||||||||||||||||||||||||||||||||||||||||||:
Db   241 QCCDLAPEAKQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   301 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRAFTEAMTRYSAPPGDLPQPEYDL 360

Qy   365 ELITSCSSNVSVAHDASGKRVVYLTRDPTTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ELITSCSSNVSVAHDASGKRVVYLTRDPATPLARAAWETARHTPVNSWLGNIIMYAPTLW 420

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
              |||||||||||||||||:||||||||||||||||||||||||:||||||||||||||||
Db   421 ARMILMTHFFSILLAQQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 480

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   481 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 540

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
              ||||||||||||||||||||||:||||
Db   541 AASRLDLSGWFVAGYSGGDIYRSLSRA 567

RESULT 11
AAG79559
ID   AAG79559 standard; protein; 545 AA.
XX
AC   AAG79559;
XX
DT   09-DEC-2002 (first entry)

DE   HCV decrease-affinity NS5B polymerase, NS5Bdelta57-HT.
XX
KW   Enzyme; hepatitis C virus; HCV; decreased affinity; NS5B; polymerase;
KW   inhibitor; RNA-dependent RNA polymerase.
XX
OS   Hepatitis C virus.
XX
PN   WO200270739-A2.
XX
PD   12-SEP-2002.
XX
PF   06-MAR-2002; 2002WO-CA000323.
XX
PR   08-MAR-2001; 2001US-0274374P.
XX
PA   (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX
PI   Kukolj G, Mckercher G;
XX
DR   WPI; 2002-698759/75.
XX
PT   Identifying inhibitors HCV NS5B RNA-dependent RNA polymerase, comprises
PT   incubating an HCV NS5B polymerase having decreased affinity for the
PT   primer template relative to the native polymerase, with a potential
PT   inhibitor.
XX
PS   Claim 10; Page 48-49; 52pp; English.
XX
CC   The sequences given in AAG79556-61 are hepatitis C virus (HCV) decreased
CC   affinity NS5B polymerases. These enzymes are used in the method of the
CC   invention for identifying a potential inhibitor of the binding between a
CC   HCV NS5B RNA-dependent RNA polymerase and an appropriate primer-template.
CC   The method comprises incubating the HCV NS5B polymerase with the primer-
CC   template in the presence and absence of a potential inhibitor. The HCV
CC   NS5B polymerase has a decreased affinity for the primer-template relative
CC   to that of native HCV NS5B RNA-dependent RNA polymerase. The method is
CC   useful for identifying a potential inhibitor of the binding between a HCV
CC   NS5B RNA-dependent RNA polymerase and a primer-template. Use of
CC   polymerase constructs having a lower affinity towards the primer-template
CC   than that of native NS5B polymerase is particularly useful for
CC   identifying potential inhibitors in screening large libraries of
CC   compounds. The new method reduces the difficulties and disadvantages of
CC   prior art. The present method provides an assay that is easy to perform
CC   on large libraries of compounds, and has improved sensitivity for
CC   detecting inhibitors that would not be identified as such using native
CC   NS5B polymerase. This protein represents a soluble form of mature HCV
CC   NS5B which lacks the C-terminal 57 amino acids and has an C-terminal hexa
CC   -histidine tag
XX
SQ   Sequence 545 AA;

Query Match   89.5%; Score 2734.5; DB 5; Length 545;
Best Local Similarity   90.1%; Pred. No. 5.1e-254;
Matches 519; Conservative 19; Mismatches 7; Indels 31; Gaps 2;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
              |||||||||||||||||||||||||||||||||||||||||||:||:||||||||||||||
Db   1 SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL 60

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
              ||||||||||||||||||||||||||||||||||||||||||||||||:|:||||||||
Db   61 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR 120

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
              |||||||||:|||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 180

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
              |||||||||||||||||:||||||||||:|||||||||||||||||||||:|||||||||
Db   181 TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCCPMGFSYDTRCFDSTVTESDIRVEESIY 240

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
              |||||||||||||:|||||:||||||||||||||||||||||||||||||||||||||||:
Db   241 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304

```
Db  241 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 300

Qy  305 AACRAAKLQDCTMLVNGDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db  301 AACRAAKLQDCTMLVNGDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL 360

Qy  365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db  361 ELITSCSSNVSVAHDASGKRVYLTRDFTTPLARAAWETARHTPINSWLGNIIMYAPTLW 420

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 480

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            ||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  481 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKL------ 534

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH 580
            :|:|-----------||||||
Db  535 -AAALE-------------HHHHHH 545
```

RESULT 12
AAR58592
ID AAR58592 standard; protein; 462 AA.
XX
AC AAR58592;
XX
DT 09-MAY-1995 (first entry)
XX
DE Hepatitis C virus 8003-9388 fragment antigen.
XX
KW Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
OS Hepatitis C virus.
XX
PN JP06225770-A.
XX
PD 16-AUG-1994.
XX
PF 08-JUL-1993; 93JP-00193104.
XX
PR 10-JUL-1992; 92JP-00207391.
XX
PA (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
PA (SANW) SANWA KAGAKU KENKYUSHO CO.
PA (TOFU) TONEN CORP.
PA (KOKU-) KOKUSAI SHIYAKU KK.
XX
DR WPI; 1994-298800/37.
DR N-PSDB; AAQ70542.
XX
PT A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for
PT diagnosis of NANBH and detection of HCV.
XX
CC Claim 1; Page 16; 22pp; Japanese.
XX
CC AAQ70542 is a fragment of hepatitis C virus (HCV) or non-A non-B hepatits
CC virus (NANBH) non-structural coding region five, encompassing base pairs
CC 8003-9388. It codes for AAR58592 an antigen to a non-structural region of
CC the HCV virus, which can be used in the diagnosis of NANBH patients and
CC the detection of HCV carriers
XX
SQ Sequence 462 AA;

Query Match   74.6%; Score 2279; DB 2; Length 462;
Best Local Similarity 93.0%; Pred. No. 3.2e-210;
Matches 426; Conservative 18; Mismatches 14; Indels 0; Gaps 0;

```
Qy  117 SRAVNHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEK 176
            |||||||||||||||:||:||||||||||||||||||||||||||||||||||||||
Db    1 SKAVNHIRSVWKDLLEDNVTPVDTTVMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVYEK 60
```

```
Db  177 MALYDVVSTLPQAVMGGSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTEND 236
            ||||:|||||||||||||||:|||||||||||||||||||||||||:|||||||||||
Qy   61 MALYNVVSTLPQAVMGGSYGLQYSPGQRVEFLVNAWKSKKCPMGFSYDNRCFDSTVTESD 120

Qy  237 IRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNT 296
            ||:|||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db  121 IRIEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNT 180

Qy  297 LTCYLKAAAACRAAKLQDCTMLVNGDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGD 356
            ||||||:|:||||||||||||||||||||||||||||||||:|||||||||||||||||
Db  181 LTCYLKASAACRAAKLRDCTMLVNGDLVVICESAGTQEDEANLRVFTEAMTRYSAPPGD 240

Qy  357 PPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNI 416
            |:|||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db  241 PPRPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPSTPLARAAWETARHTPVNSWLGNI 300

Qy  417 IMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 476
            |:|||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db  301 IMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 360

Qy  477 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRT 536
            ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db  361 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRT 420

Qy  537 KLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR 574
            |||||||||||:|:||||||:||||||||||||||||
Db  421 KLKLTPIPAASQLDLSSWFLAGYSGEHIYHSLSRARPR 458
```

RESULT 13
ADL17783
ID ADL17783 standard; protein; 462 AA.
XX
AC ADL17783;
XX
DT 06-MAY-2004 (first entry)
XX
DE Hepatitis C virus protein (clone C6-82).
XX
KW non-A-non-B type hepatitis virus antigen;
KW hepatitis C virus antibody detection; clone C260-1.
XX
OS Hepatitis C virus.
XX
PN JP2004000150-A.
XX
PD 08-JAN-2004.
XX
PF 24-FEB-2003; 2003JP-00046379.
XX
PR 10-JUL-1992; 92JP-00207391.
PR 08-JUL-1993; 93JP-00193104.
XX
PA (KOKU-) KOKUSAI SHIYAKU KK.
XX
DR WPI; 2004-085213/09.
DR N-PSDB; ADL17778.
XX
PT Novel non-A-non-B type hepatitis virus nucleic acid fragment useful for
PT diagnosing non-A-non-B type
PT hepatitis patients, as well as for detecting those people who are
PT virus carrier.
XX
PS Disclosure; Page 32-33; 35pp; Japanese.
XX
CC This invention relates to a novel isolated nucleic acid fragment encoding
CC a non-A-non-B type hepatitis virus antigen protein sequence.
CC Specifically, it refers to a method for diagnosing non-A-non-B type
CC hepatitis patients. The present invention describes a diagnosis method that uses a
CC hepatitis C virus antibody detection reagent and subsequent PCR
CC amplification to identify a clone C260-1 that encodes the non-A-non-B
CC type hepatitis virus antigen sequence. This polypeptide sequence is the .

```
CC   non-A-non-B type hepatitis virus protein fragment of the invention. NOTE:
CC   This sequence is given in the sequence listing as an embedded protien and
CC   is not further referred to in the specification.
XX
SQ   Sequence 462 AA;

Query Match             74.6%;  Score 2279;  DB 8;  Length 462;
Best Local Similarity   93.0%;  Pred. No. 3.2e-210;
Matches  426;  Conservative  18;  Mismatches  14;  Indels  0;  Gaps  0;

Qy  117 SRAVNHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEK 176
        |||:|||||||||||||  :|||:|||||||||||||||||||||||||||||||||:|
Db    1 SKAVNHIRSVWKDLLEDNVTPVDTTVMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVYEK 60

Qy  177 MALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTEND 236
        |||:||:||||||||||||||:||||:|||||||:|:|||:|||||||:|||||||||:|
Db   61 MALYNVVSTLPQAVMGSSYGLQYSPGQRVEFLVNAWKSKKCPMGFSYDNRCFDSTVTESD 120

Qy  237 IRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNT 296
        ||:|||||||||||||||||:||||||||:||||||||||||||||||||||||||||||
Db  121 IRIEESIYQCCDLAPEARQAIKSLTERLYIIGGPLTNSKGQNCGYRRCRASGVLTTSCGNT 180

Qy  297 LTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGD 356
        |||||||:|||||||:|||||||||||||||||||||||||:||||||||||||||||||
Db  181 LTCYLKASAACRAAKLRDCTMLVNGDDLVVICESAGTQEDEANLRVFTEAMTRYSAPPGD 240

Qy  357 PQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNI 416
        |:|||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db  241 PPRPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPSTPLARAAWETARHTPVNSWLGNI 300

Qy  417 IMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 476
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 IMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF 360

Qy  477 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRT 536
        |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db  361 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRT 420

Qy  537 KLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR 574
        ||||||||||| :|:||||| || |||:||||||||||
Db  421 KLKLTPIPAASQLDLSSWFLAGYSGEHIYHSLSRARPR 458

RESULT 14
AAR67630
ID   AAR67630 standard; protein; 547 AA.
XX
AC   AAR67630;
XX
DT   27-AUG-2003  (revised)
DT   25-MAR-2003  (revised)
DT   18-AUG-1995  (first entry)
XX
DE   Non-A Non-B hepatitis virus non-structural protein.
XX
KW   Non-A Non-B hepatitis virus; structural region; cDNA to genomic RNA;
KW   detection; reagent; anti-Non-A Non-B hepatitis virus antibody; vaccine;
KW   antigen; epitope; diagnosis.
XX
OS   Non-A,
OS   non-B hepatitis virus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         52..72
FT                   /note= "putative epitope site (see AAR67641)"
XX
PN   EP628572-A2.
XX
PD   14-DEC-1994.
XX
PF   27-MAY-1994;  94EP-00108256.
XX
PR   28-MAY-1993;  93JP-00126709.
```

```
PR   02-MAR-1994;  94JP-00032201.
XX
PA   (EISA ) EISAI CO LTD.
XX
PI   Aoyama M, Obara T, Tohmatsu J, Sawada T, Hosoda T, Iwasaki Y;
PI   Arima T;
XX
DR   WPI; 1995-015655/03.
DR   N-PSDB; AAQ75819.
XX
PT   New non-A non-B hepatitis virus sub-type - used to develop prods. for
PT   detection, diagnosis, prevention and treatment of non-A non-B hepatitis.
XX
PS   Claim 1; Page 33-37; 59pp; English.
XX
CC   AAR67630 is the structural protein of the Non-A Non-B (NANB) hepatitis
CC   virus encoded by a partial cDNA to genomic RNA sequence. Regions of the
CC   polypeptide were studied for suitability as an epitope. The novel
CC   epitopes are effective in the diagnosis of NANB hepatitis. The nucleotide
CC   sequences (see AAQ75817-19) were isolated from the plasma of donors in
CC   Japan with high s-GTP levels, and were found to be different from
CC   previously reported NANB hepatitis viruses. The DNA can be used as a
CC   reagent for detecting the NANB hepatitis viral gene. The polypeptides can
CC   be used as reagents for detecting anti-NANB hepatitis antibodies or as a
CC   NANB hepatitis viral vaccine. (Updated on 25-MAR-2003 to correct PN
CC   field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 547 AA;

Query Match             70.2%;  Score 2144;  DB 2;  Length 547;
Best Local Similarity   75.5%;  Pred. No. 4.4e-197;
Matches  397;  Conservative  55;  Mismatches  74;  Indels  0;  Gaps  0;

Qy   49 ASLRQKKVTFDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVBEACKLTPPHSAKSKFGYG 108
        |||| |||||||:|||| ||| :|||:||:: |:::|:|||||:||:||||||||||:|:|
Db    1 ASLRAKKVTFDRVQVLDAHYDSVLQDVKRAASKVSARLLSVBEACALTPPHSAKSLYGFG 60

Qy  109 AKDVRSLSSRAVNHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPD 168
        ||:||||| |||||||||||:|||||:||||:|||||||||||:| |||||||||||||||
Db   61 AKEVRSLSRRAVNHIRSVWEDLLEDQHTPINTTIMAKNEVFCIDPAKGGRKPARLIVFPD 120

Qy  169 LGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCF 228
        |||||||||||||: ::||: ||| ||||||||| :||:|| |:| ||| |||||||||||
Db  121 LGVRVCEKMALYDIAQKLPKAIMGPSYGFQYSPAERVDFLLKAWGSKKDPMGFSYDTRCF 180

Qy  229 DSTVTENDIRVEBESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGV 288
        ||||||.||| ||||||| | | |||||: |:||||||||||||||||| ||||||||||||
Db  181 DSTVTERDIRTESIYQACSLPQEARTVIHSLTERLYVGGPMTNSKGQSCGYRRCRASGV 240

Qy  289 LTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMT 348
        :||| |||:|||:||| ||:|| : | ||| ||||||| ||| |:|||:  |||:||||||
Db  241 FTTSMGNTMTCYIKALAACKAAGIKDPIMLVCGDDLVVISESQGNEEDERNLRAFTEAMT 300

Qy  349 RYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTP 408
        |||||||||| ||||||||||||||||||  | |::|||||||||.|:||||||||.||.|
Db  301 RYSAPPGDPPRPEYDLELITSCSSNVSVALDPRGRRRYYLTRDPTTPISRAAWETVRHSP 360

Qy  409 VNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIE 468
        ||||||||| |||| |.||::|||||: |||| : |::.|:::: |.:||:  |||| ||E
Db  361 VNSWLGNIIQYAPTIWVRMVIMTHFFAILLAQDTLNQNINFEMYGAVYSVNPLDLPAIIE 420

Qy  469 RLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKY 528
        ||||| |||||:|||  :.|||| | |||||||| ||  || ||| || |:|||||||: |
Db  421 RLHGLDAFSLHTYSPHELSRVAATLRKLGAPPLRAWKSRARAVRASLIAQGGRAAICGRY 480

Qy  529 LFNWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR 574
        ||||||:|||||||:| |||||||||| || :|||||:|:||||||
Db  481 LFNWAVKTKLKLTPLPEAARLDLSGWFTVGAGGGDIFHSVSRARPR 526

RESULT 15
AAR29906
ID   AAR29906 standard; protein; 389 AA.
```

```
Qy  340  LRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARA  399
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db  121  LRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARA  180

Qy  400  AWETARHTPVNSWLGNIIMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIE   459
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  181  AWETARHTPVNSWLGNIIMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGATYSIE   240

Qy  460  PLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQG  519
         |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  PLDLPQIIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQG  300

Qy  520  GRAAICGKYLFNWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR  574
         |||| ||||||||||:|||||||||:|:||||||||||||||||||||||||||
Db  301  GRAATCGKYLFNWAVKTKLKLTPIPEASQLDLSGWFVAGYSGGDIYHSLSRARPR  355
```

Search completed: September 22, 2005, 14:51:00
Job time : 86 secs

```
XX
AC   AAR29906;
XX
DT   25-MAR-2003  (revised)
DT   26-APR-1993  (first entry)
XX
DE   HCV NS4-NS5 peptide O30-4.
XX
KW   Clone; polypeptide; NS4-NS5; Hepatitis C; virus; HCV; serum; HC;
KW   transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
KW   turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.
XX
OS   Hepatitis C virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 377
FT                   /note= "Nonsense mutation"
FT   Misc-difference 381
FT                   /note= "Nonsense mutation"
XX
PN   EP518313-A2.
XX
PD   16-DEC-1992.
XX
PF   11-JUN-1992;  92EP-00109812.
XX
PR   11-JUN-1991;  91JP-00139268.
PR   12-JUL-1991;  91JP-00172794.
PR   07-OCT-1991;  91JP-00287008.
PR   16-DEC-1991;  91JP-00332329.
PR   20-APR-1992;  92JP-00099957.
XX
PA   (MITU ) MITSUBISHI KASEI CORP.
XX
PI   Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
XX
DR   WPI; 1992-417213/51.
DR   N-PSDB; AAQ32537.
XX
PT   New hepatitis C virus gene and its encoded protein - used for diagnosing
PT   and vaccinating against hepatitis C virus infections.
XX
PS   Disclosure; Page 257-59; 305pp; English.
XX
CC   The sequences given in AAR29871-906 and AAR29533 are encoded by various
CC   clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
CC   invention. These NS4-NS5 RNA sequences were isolated from the serum of a
CC   patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC   converted into cDNA using transcriptase in the presence of one of the
CC   primer sequences given in AAQ32565-77. The sequences were then amplified
CC   using primer pairs. The cDNA sequences isolated represent different
CC   alleles of the same region of the HCV gene. Sequence analysis shows that
CC   these clones represent the core region and some upstream sequences of
CC   HCV. These polypeptides are thought to contain a highly hydrophilic
CC   region which can adopt a "turn structure" which is not an alpha helix or
CC   a beta sheet. These polypeptides are thought to act as antigen
CC   determinants and are highly reactive with antiserum raised against HCV-
CC   associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 389 AA;
```

Query Match   59.3%; Score 1812; DB 2; Length 389;
Best Local Similarity 96.1%; Pred. No. 2.6e-165;
Matches 341; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```
Qy  220  GFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCG  279
         | ||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||
Db    1  GIPYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCG   60

Qy  280  YRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAS  339
         |||||||||||||||||||||||||:|||||||||||||||:|||||||||||||||||||
Db   61  YRRCRASGVLTTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAAS  120
```

This Page Blank (uspto)

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:    September 22, 2005, 14:47:55 ; Search time 43 Seconds
                                          (without alignments)
                                          1297.808 Million cell updates/sec

Title:          US-10-712-479-4
Perfect score:  3055
Sequence:       1 MASMSMSYTWTGALITPCAA.........GGDIYHSLSRARPRHHHHHH 580

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:   283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:    248019

Minimum DB seq length: 0
Maximum DB seq length: 585

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR 79:*
          1: pir1:*
          2: pir2:*
          3: pir3:*
          4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1777	58.2	365	2	JQ0879	NS5 protein - hepa
2	1655	54.2	365	2	JQ0880	NS5 protein - hepa
3	951	31.1	189	2	S60587	non-structural pro
4	774	25.3	284	2	PS0104	genome polyprotein
5	744	24.4	259	2	PS0102	genome polyprotein
6	667	21.8	135	2	S44214	genome polyprotein
7	630	20.6	156	2	D39109	genome polyprotein
8	569	18.6	113	2	PC1278	NS5 protein - hepa
9	563	18.4	113	2	PC1277	NS5 protein - hepa
10	562	18.4	113	2	PC1274	NS5 protein - hepa
11	562	18.4	113	2	PC1275	NS5 protein - hepa
12	560	18.3	113	2	PC1276	NS5 protein - hepa
13	513	16.8	113	2	PC1279	NS5 protein - hepa
14	426	13.9	113	2	PC1282	NS5 protein - hepa
15	426	13.9	113	2	PQ0275	polyprotein - hepa
16	425	13.9	113	2	PQ0276	polyprotein - hepa
17	424	13.9	113	2	PC1281	NS5 protein - hepa
18	422	13.8	113	2	PC1283	NS5 protein - hepa
19	409	13.4	113	2	PC1280	NS5 protein - hepa
20	408	13.4	113	2	PQ0273	polyprotein - hepa
21	406	13.3	113	2	PQ0274	polyprotein - hepa
22	402	13.2	78	2	PS0085	hypothetical prote
23	336	11.0	266	2	PQ0393	genome polyprotein
24	300	9.8	132	2	PQ0396	genome polyprotein
25	298	9.8	132	2	PQ0394	genome polyprotein
26	296	9.7	60	2	S44215	genome polyprotein
27	264	8.6	71	2	PC1300	genome polyprotein
28	261	8.5	71	2	PC1302	genome polyprotein
29	260	8.5	71	2	PC1301	genome polyprotein
30	225	7.4	52	2	PQ0555	nonstructural prot
31	224	7.3	52	2	PQ0558	nonstructural prot
32	223	7.3	52	2	PQ0556	nonstructural prot
33	221	7.2	52	2	PQ0554	nonstructural prot
34	215	7.0	52	2	PQ0557	nonstructural prot
35	131	4.3	535	2	B49600	probable polymeras
36	128.5	4.2	532	1	RRVQC2	RNA-directed RNA p
37	123.5	4.0	581	2	JQ1879	hypothetical 65.4K
38	106	3.5	317	2	AC3410	peptidylprolyl iso
39	105.5	3.5	356	2	G96806	thaumatin-like pro
40	103	3.4	315	2	T16359	hypothetical prote
41	96.5	3.2	546	2	B40407	sterol carrier pro
42	94	3.1	315	2	AB3202	hypothetical prote
43	92.5	3.0	565	2	T47423	hypothetical prote
44	91	3.0	523	2	JC7556	linoleoyl-CoA desa
45	90	2.9	476	2	AE2126	pyruvate kinase [i

ALIGNMENTS

```
RESULT 1
JQ0879
NS5 protein - hepatitis C virus (strain J4) (fragment)
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0879
R;Okamoto, H.
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0879
A;Accession: JQ0879
A;Molecule type: genomic RNA
A;Residues: 1-365 <OKA>
A;Cross-references: UNIPROT:Q81717
A;Experimental source: strain J4
C;Superfamily: hepatitis C virus genome polyprotein

Query Match          58.2%;   Score 1777;  DB 2;  Length 365;
Best Local Similarity  96.8%;   Pred. No. 7e-129;
Matches  333;  Mismatches  8;  Conservative  3;  Indels  0;  Gaps  0;

Qy  231  TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLT  290
         ||||:||||||||||||||||||||||||||||||:|||:||||||||||||||||||||
Db    1  TVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLT  60

Qy  291  TSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRY  350
         |||||||||||||:||||||||||||||||||||||||||||||||||:||||||||||||
Db   61  TSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRVFTEAMTRY  120

Qy  351  SAPPGDPPQPEYDLELITSCSSSNVSVAHDASGKRVYYLTRDPVPVPLARAAWETARHTPVN  410
         ||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||||||
Db  121  SAPPGDPPQPEYDLELITSCSSSNVSVAHDASGKRVYYLTRDPTIPLARAAWETARHTPVN  180

Qy  411  SWLGNIIMYAPTIWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERL  470
         |||||||||:||||||||||||||||||||||:|||||||||||||||||||||||||||
Db  181  SWLGNIIMYAPALWARMILMTHFFSILLAQEQEQALDCQIYGACYSIEPLDLPQIIERL  240

Qy  471  HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLF  530
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||:||
Db  241  HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLF  300

Qy  531  NWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR  574
         |||||||||||||||||:||||||||||||||||||||||||||
Db  301  NWAVRTKLKLTPIPAASQLDLSGWFVAGYSGGDIYHSLSRARPR  344

RESULT 2
JQ0880
NS5 protein - hepatitis C virus (strain J1) (fragment)
C;Species: hepatitis C virus
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: JQ0880
```

```
R;Okamoto, H.
submitted to JIPID, January 1991
A;Description: The 5'-terminal and 3'-terminal sequences of the genomic RNA of hepatitis
A;Reference number: JQ0879
A;Accession: JQ0880
A;Molecule type: genomic RNA
A;Residues: 1-365 <OKA>
A;Cross-references: UNIPROT:Q81716
A;Experimental source: strain J1
C;Superfamily: hepatitis C virus genome polyprotein

Query Match           54.2%;  Score 1655;  DB 2;  Length 365;
Best Local Similarity 88.7%;  Pred. No. 1.7e-119;
Matches 305;  Conservative 22;  Mismatches 17;  Indels 0;  Gaps 0;

Qy  231 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLT 290
            |||||||||||||||||||:|||||||||||||||||||::||||||||||||||||||
Db    1 TVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLT 60

Qy  291 TSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRY 350
            ||||||||||||   ||||||||||||||||||||||||| : |||||||: :|||||||
Db   61 TSCGNTLTCYIKAQAACRAAGLRDCTMLVCGDDLVVICESQGVQEDAASLRAFTEAMTRY 120

Qy  351 SAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVN 410
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  121 SAPPGDPPQPEYDLELITSCSSNVSVAHDGTGKRVYLTRDPTTPLARAAWETARHTPVN 180

Qy  411 SWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERL 470
            ||||||||:|||||||||||||||||:||||::|||||||:|||||||||||||:|||||
Db  181 SWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACYSIEPLDLPPIIQRL 240

Qy  471 HGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGQRAAICGKYLF 530
            ||||||||||||||||||||:|||||||||||:|||||||||||:||:|:||||||||||
Db  241 HGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLF 300

Qy  531 NWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPR 574
            |||||||||||||:||:||||||||:|||||||||||:|:||||
Db  301 NWAVRTKLKLTPIAAAGRLDLSGWFTAGYSGGDIYHSVSHARPR 344

RESULT 3
S60587
non-structural protein NS5 - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C;Accession: S60587
R;Xu, L.Z.; Martinot-Peignoux, M.; Marcellin, P.; Benhamou, J.P.; Larzul, D.
J. Hepatol. 20, 598-602, 1994
A;Title: Comparison of the sensitivity of nested PCR in the 5' non-coding and the NS5 re
A;Reference number: S60587; MUID:94351175; PMID:8071535
A;Status: preliminary; translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-189 <XUL>
A;Cross-references: UNIPROT:Q68959; EMBL:X59934; NID:g556891; PIDN:CAA42557.1; PID:g4388
C;Superfamily: hepatitis C virus genome polyprotein

Query Match           31.1%;  Score 951;  DB 2;  Length 189;
Best Local Similarity 93.7%;  Pred. No. 9.4e-66;
Matches 177;  Conservative 5;  Mismatches 7;  Indels 0;  Gaps 0;

Qy  153 PEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTW 212
            |||||||||||||:|||||||||||||||||||||||||||||||||||:|||||||:|
Db    1 PEKGGRKPARLIVYPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAW 60

Qy  213 KAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTN 272
            |:|||||||:||||||||||||:||||||||||||||||||||:|||||||||:|||::|
Db   61 KSKKCPMGFAYDTRCFDSTVTESDIRVEESIYQCCDLAPEARQVIRSLTERLYIGGPLTN 120

Qy  273 SKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAG 332
            ||||||||||||||||||||||||||||||:||||||||||||||||:||||||||||:|
Db  121 SKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESGG 180
```

```
Qy  333 TQEDAASLR 341
            |||||||||
Db  181 TQEDAASLR 189

RESULT 4
PS0104
genome polyprotein - hepatitis C virus (isolate pt) (fragments)
N;Contains: NS5 protein; structural protein
C;Species: hepatitis C virus
A;Note: host Pan troglodytes (chimpanzee)
C;Accession: PS0104
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
R;Takeuchi, K.; Boonmar, S.; Kubo, Y.; Katayama, T.; Harada, H.; Ohbayashi, A.; Choo, Q.I
Gene 91, 287-291, 1990
A;Title: Hepatitis C viral cDNA clones isolated from a healthy carrier donor implicated j
A;Reference number: PS0102; MUID:91007289; PMID:2170237
A;Accession: PS0104
A;Molecule type: genomic RNA
A;Residues: 1-284 <TAK>
A;Cross-references: UNIPROT:Q81633; GB:D90078; GB:D90079
A;Note: the authors translated the codon TGT for residue 274 as Lys
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
F;1-129/Product: structural protein (fragment) #status predicted <XPR>
F;130-284/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match           25.3%;  Score 774;  DB 2;  Length 284;
Best Local Similarity 88.6%;  Pred. No. 6.6e-52;
Matches 140;  Conservative 13;  Mismatches 5;  Indels 0;  Gaps 0;

Qy  378 HDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWARMLMTHFFSIL 437
            :|| |||||||||| ||||||||||||||||||||||:||||||||||||||||:|:|
Db  127 HKPAGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLWARMLMTHFPSVL 186

Qy  438 LAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLG 497
            |||||||||||:|||||||||||||||:|||:|||||||||||||||||||||:||||||
Db  187 IARDQLEQALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPGEINRVAACLRKLG 246

Qy  498 VPPLRVWRHRARSVRAKLLSQGQRAAICGKYLFNWAVR 535
            ||||| ||||||||||:|| ::|||||||||||||||
Db  247 VPPLRAWRHRARSVRARLLARGGRAAICGKYLFNWAVR 284

RESULT 5
PS0102
genome polyprotein - hepatitis C virus (isolate J1) (fragments)
N;Contains: NS5 protein; structural protein
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Accession: PS0102
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
R;Takeuchi, K.; Boonmar, S.; Kubo, Y.; Katayama, T.; Harada, H.; Ohbayashi, A.; Choo, Q.I
Gene 91, 287-291, 1990
A;Title: Hepatitis C viral cDNA clones isolated from a healthy carrier donor implicated j
A;Reference number: PS0102; MUID:91007289; PMID:2170237
A;Contents: J1
A;Accession: PS0102
A;Molecule type: genomic RNA
A;Residues: 1-259 <XUL>
A;Note: the authors translated the codon TGT for residue 255 as Lys
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
F;1-117/Product: structural protein (fragment) #status predicted <SPR>
F;118-259/Product: NS5 protein (fragment) #status predicted <NS5>

Query Match           24.4%;  Score 744;  DB 2;  Length 259;
Best Local Similarity 99.3%;  Pred. No. 1.2e-49;
Matches 141;  Conservative 0;  Mismatches 0;  Indels 1;  Gaps 0;

Qy  388 LTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWARMLMTHFFSILLAQEQLEKAL 447
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   118 LTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKAL 177
             ||||||||||::||:::|||||||||||||||||||||||||||||||||||||||||||||
Db   66 RHRARSVRARLLARGGRAAICGKYLFNWAVRTKLKLTPIAAAGQDLSGWFTAGYSGGDI 125

Qy   448 DCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHR 507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   565 YHSLSRARPR 574
Db   178 DCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHR 237
             |||||:| |||
Db   126 YHSVSHARPR 135

Qy   508 ARSVRAKLLSQGGRAAICGKYL 529
             |||||||||||||||||||||||
Db   238 ARSVRAKLLSQGGRAATCGKYL 259

RESULT 6
S44214
genome polyprotein - hepatitis C virus (fragment)
N;Contains: NS5a protein
C;Species: hepatitis C virus
C;Date: 13-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S44214
R;Feucht, H.H.
submitted to the EMBL Data Library, April 1994
A;Reference number: S44213
A;Accession: S44214
A;Molecule type: mRNA
A;Residues: 1-135 <FEU>
A;Cross-references: UNIPROT:Q68962; EMBL:X78954; NID:g9475176; PIDN:CAA55551.1; PID:g9398
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
F;1-135/Product: NS5a protein #status predicted <MAT>

Query Match   21.8%;   Score 667; DB 2; Length 135;
Best Local Similarity 91.9%;   Pred. No. 4.1e-44;
Matches 124; Conservative   6; Mismatches   5; Indels   0; Gaps   0;

Qy   147 EVFCVQPEKGGRKPARLIVEFDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVE 206
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1 EVFCVQPEKGGRKPARLIVEFDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVE 60

Qy   207 FLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYV 266
             ||||| ||||||||||||||||||||||||||||||||||||||||||| :||||||
Db   61 FLVNAWKSKKNPMGFAYCTRCFDSTVTESDIRVEESIYQCCDLAPQARQAIRLLTERLYI 120

Qy   267 GGPMTNSKGQNCGYR 281
             |||:||||||||||
Db   121 GGPLTNSKGQNCGYR 135

RESULT 7
D39109
genome polyprotein - hepatitis C virus (fragments)
C;Species: hepatitis C virus
C;Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 17-Nov-2000
C;Accession: D39109
R;Han, J.H.; Shyamala, V.; Richman, K.H.; Brauer, M.J.; Irvine, B.; Urdea, M.S.; Tekamp-
Proc. Natl. Acad. Sci. U.S.A. 88, 1711-1715, 1991
A;Title: Characterization of the terminal regions of hepatitis C viral RNA: identificati
A;Reference number: A39109; MUID:91156678; PMID:1705704
A;Accession: D39109
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-156 <HAN>
A;Cross-references: GB:M58406; GB:M58407
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

Query Match   20.6%;   Score 630; DB 2; Length 156;
Best Local Similarity 89.2%;   Pred. No. 3.5e-41;
Matches 116; Conservative   8; Mismatches   8; Indels   6; Gaps   0;

Qy   445 KALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVW 504
             |||| ::||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db   6 KALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAW 65

Qy   505 RHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDI 564

Db   66 RHRARSVRARLLARGGRAAICGKYLFNWAVRTKLKLTPIAAAGQDLSGWFTAGYSGGDI 125

Qy   565 YHSLSRARPR 574
             |||:| ||||
Db   126 YHSVSHARPR 135

RESULT 8
PC1278
NS5 protein - hepatitis C virus (strain K1-4) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1278
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1278
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01187; GB:D10645; GB:D90554; NID:g221668; PIDN:BAA01492.1; I
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match   18.6%;   Score 569; DB 2; Length 113;
Best Local Similarity 95.6%;   Pred. No. 1.1e-36;
Matches 108; Conservative   3; Mismatches   2; Indels   0; Gaps   0;

Qy   230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 STVTENDIRVEESIYQCCDLAPEARQVIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy   290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
             |||||||||||||:||| ||||||||||||| ||||||||||||||||||||
Db   61 TTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRV 113

RESULT 9
PC1277
NS5 protein - hepatitis C virus (strain K1-3) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1277
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1277
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:O70640; GB:D10644; GB:D90553; NID:g221666; PIDN:BAA01491.1; I
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match   18.4%;   Score 563; DB 2; Length 113;
Best Local Similarity 94.7%;   Pred. No. 3.2e-36;
Matches 107; Conservative   4; Mismatches   2; Indels   0; Gaps   0;

Qy   230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
             ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db   1 STVTENDIRVEESIYQCCDLAPEARQVIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy   290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
             |||||||||||||:|| |||||||||||||| ||||||||||||||||||||
Db   61 TTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRV 113

RESULT 10

```
PC1274
NS5 protein - hepatitis C virus (strain K1) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1274
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1274
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01192; GB:D10641; GB:D90550; NID:g221658; PIDN:BAA01488.1;
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match             18.4%;     Score 562;  DB 2;  Length 113;
Best Local Similarity   95.6%;          Pred. No. 3.8e-36;
Matches  108;   Conservative   2;   Mismatches   2;   Indels   3;   Gaps   0;

Qy     230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
           |||||||||||||||| ||||||||||||||||||:|||||:|||||||||||||||||||
Db       1 STVTENDIRVEESIYQSCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy     290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
           ||||||||||||| ||||||||||||||||| |||||||||||||||||||||
Db      61 TTSCGNTLTCYLKATAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRV 113

RESULT 11
PC1275
NS5 protein - hepatitis C virus (strain K1-1) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1275
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1275
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01184; GB:D10642; GB:D90551; NID:g221662; PIDN:BAA01489.1;
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match             18.4%;     Score 562;  DB 2;  Length 113;
Best Local Similarity   94.7%;          Pred. No. 3.8e-36;
Matches  107;   Conservative   4;   Mismatches   2;   Indels   2;   Gaps   0;

Qy     230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
           |||||||||||||||| |||||||||||||||||:|||||:|||||||||||||||||||
Db       1 STVTENDIRVEESIYQSCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy     290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
           ||||||||||||| ||||||||||||||| ||||||||||||||||||| |||
Db      61 TTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAANLRV 113

RESULT 12
PC1276
NS5 protein - hepatitis C virus (strain K1-2) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1276
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;
```

```
A;Accession: PC1276
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01185; GB:D10643; GB:D90552; NID:g221664; PIDN:BAA01490.1; I
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match             18.3%;     Score 560;  DB 2;  Length 113;
Best Local Similarity   93.8%;          Pred. No. 5.4e-36;
Matches  106;   Conservative   5;   Mismatches   2;   Indels   0;   Gaps   0;

Qy     230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
           |||||:||||||||||||||||||||:|||||||||:|||||:|||||||||||||||||
Db       1 STVTESDIRVEESIYQCCDLAPEARQVIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy     290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
           ||||||||||||| ||||||||||||||| ||||||||||||||||||||||
Db      61 TTSCGNTLTCYLKASAACRAAKLRDCTMLVCGDDLVVICESAGTQEDAASLRV 113

RESULT 13
PC1279
NS5 protein - hepatitis C virus (strain PT-1) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1279
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1279
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01193; GB:D10646; GB:D90555; NID:g221670; PIDN:BAA01493.1; I
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein

Query Match             16.8%;     Score 513;  DB 2;  Length 113;
Best Local Similarity   85.7%;          Pred. No. 2.2e-32;
Matches   96;   Conservative   9;   Mismatches   9;   Indels   7;   Gaps   0;

Qy     230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
           |||||:|||:::|||||||||:|| |:|||||||||||||||:|||:|:|||||||||||
Db       1 STVTESDIRTEBAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVL 60

Qy     290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLR 341
           |||||||||||:||:||||||:|:|||||| ||||||||||||:||||||||
Db      61 TTSCGNTLTCYIKARAACRAAGLRDCTMLVCGDDLVVICESAGVQEDAASLR 112

RESULT 14
PC1282
NS5 protein - hepatitis C virus (strain K2b) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1282
R;Enomoto, N.; Tanaka, A.; Nakao, T.; Date, T.
Biochem. Biophys. Res. Commun. 170, 1021-1025, 1990
A;Title: There are two major types of hepatitis C virus in Japan.
A;Reference number: PC1274; MUID:90358793; PMID:2117923
A;Accession: PC1282
A;Molecule type: mRNA
A;Residues: 1-113 <ENO>
A;Cross-references: UNIPROT:Q01189; GB:D10649; GB:D90558; NID:g221676; PIDN:BAA01496.1; I
C;Genetics:
A;Gene: NS5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein
```

```
Query Match           13.9%;   Score 426; DB 2;   Length 113;
Best Local Similarity  75.0%;   Pred. No. 1.1e-25;
Matches  84;   Conservative   7;   Mismatches  21;   Indels   0;   Gaps   0;

Qy      230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
                |||||| |||  |||||| |    ||| ||    ||||||||||||   |||||||||||
Db        1 STVTERDIRTEESIYQACSLPQEARTAIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVF  60

Qy      290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLR 341
                ||||||:|||:|||:||: ||  ||||| ||||||||  : || ||   :||
Db       61 TTSMGNTMTCYIKALAACKAAGIVDPIMLVCGDDLVVISESEGNEEDERNLR 112


RESULT 15
PQ0275
polyprotein - hepatitis C virus (isolate T-9) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: PQ0275
R;Mori, S.; Kato, N.; Yagyu, A.; Tanaka, T.; Ikeda, Y.; Petchclai, B.; Chiewsilp, P.; Ku
Biochem. Biophys. Res. Commun. 183, 334-342, 1992
A;Title: A new type of hepatitis C virus in patients in Thailand.
A;Reference number: PQ0273; MUID:92181465; PMID:1311926
A;Accession: PQ0275
A;Molecule type: genomic RNA
A;Residues: 1-113 <MOR>
A;Cross-references: UNIPROT:Q81819; DDBJ:D10080; NID:g221684; PIDN:BAA00974.1; PID:g2221€
A;Note: this sequence corresponds to NS5 protein
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

Query Match           13.9%;   Score 426; DB 2;   Length 113;
Best Local Similarity  73.2%;   Pred. No. 1.1e-25;
Matches  82;   Conservative  11;   Mismatches  19;   Indels   0;   Gaps   0;

Qy      230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
                |||||| || :||||||:|||||:||: :|  |||||||||||    |||||||||||
Db        1 STVTEHDIRTEEIYQCCDLEPEARKAISALTERLYIGGPMYNSKGLQCGYRRCRASGVL  60

Qy      290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLR 341
                ||||||:|||:|||:||: : ||  || : | |: || : |:|:|:|
Db       61 PTSFGNTITCYIKATAASRAAGLKDPSFLVCGDDLVVVSESCGVEEDRAALR 112


Search completed: September 22, 2005, 14:54:47
Job time : 44 secs
```

This Page Blank (uspto)

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         September 22, 2005, 14:47:32 ; Search time 175 Seconds
                                   (without alignments)
                                   1697.177 Million cell updates/sec

Title:          US-10-712-479-4
Perfect score:  3055
Sequence:       1 MASMSMSYTWTGALITPCAA.........GGDIYHSLSRARPRHHHHHH 580

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:    1438883

Minimum DB seq length: 0
Maximum DB seq length: 585

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      UniProt 03:*
                1: uniprot_sprot:*
                2: uniprot_trembl:*

                                   SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2873	94.0	571	2	Q6B4Q1	Q6b4q1 hepatitis c
2	2870	93.9	571	2	Q66NA8	Q66na8 hepatitis c
3	2857	93.5	571	2	Q66N95	Q66n95 hepatitis c
4	2857	93.5	571	2	Q66NA6	Q66na6 hepatitis c
5	2856	93.5	571	2	Q66N84	Q66n84 hepatitis c
6	2853	93.4	571	2	Q6B4Q2	Q6b4q2 hepatitis c
7	2851	93.3	571	2	Q66N79	Q66n79 hepatitis c
8	2846	93.2	571	2	Q66N83	Q66n83 hepatitis c
9	2839	92.9	571	2	Q66NA2	Q66na2 hepatitis c
10	2837	92.9	571	2	Q66NA0	Q66na0 hepatitis c
11	2837	92.9	571	2	Q66N74	Q66n74 hepatitis c
12	2836	92.8	571	2	Q66N89	Q66n89 hepatitis c
13	2820	92.3	571	2	Q66NB0	Q66nb0 hepatitis c
14	2803	91.8	571	2	Q66NA7	Q66na7 hepatitis c
15	2679	87.7	571	2	Q66N75	Q66n75 hepatitis c
16	2673	87.5	571	2	Q66NA5	Q66na5 hepatitis c
17	2673	87.5	571	2	Q66N86	Q66n86 hepatitis c
18	2672	87.5	571	2	Q66N93	Q66n93 hepatitis c
19	2672	87.5	571	2	Q66N4	Q66n4 hepatitis c
20	2671	87.4	571	2	Q66N99	Q66n99 hepatitis c
21	2669	87.4	571	2	Q66N91	Q66n91 hepatitis c
22	2668	87.3	571	2	Q66NB3	Q66nb3 hepatitis c
23	2665	87.2	571	2	Q66N94	Q66n94 hepatitis c
24	2664	87.2	571	2	Q66N98	Q66n98 hepatitis c
25	2663	87.2	571	2	Q66NA3	Q66na3 hepatitis c
26	2661	87.1	571	2	Q66N82	Q66n82 hepatitis c
27	2661	87.1	571	2	Q66NA1	Q66na1 hepatitis c
28	2661	87.1	571	2	Q66NB5	Q66nb5 hepatitis c
29	2660	87.1	571	2	Q66N85	Q66n85 hepatitis c
30	2660	87.1	571	2	Q66N90	Q66n90 hepatitis c
31	2660	87.1	571	2	Q66N96	Q66n96 hepatitis c
32	2660	87.1	571	2	Q66NA9	Q66na9 hepatitis c
33	2660	87.1	571	2	Q66NB1	Q66nb1 hepatitis c
34	2659	87.0	571	2	Q66NB0	Q66nb0 hepatitis c
35	2658	87.0	571	2	Q66NB2	Q66nb2 hepatitis c
36	2657	87.0	571	2	Q66NB8	Q66nb8 hepatitis c
37	2656	86.9	571	2	Q66N76	Q66n76 hepatitis c
38	2655	86.9	571	2	Q66N87	Q66n87 hepatitis c
39	2655	86.9	571	2	Q66N77	Q66n77 hepatitis c
40	2650	86.7	571	2	Q66N73	Q66n73 hepatitis c
41	2646	86.6	571	2	Q66N92	Q66n92 hepatitis c
42	2646	86.6	571	2	Q66NB1	Q66nb1 hepatitis c
43	2644	86.5	571	2	Q66NB4	Q66nb4 hepatitis c
44	2644	86.5	571	2	Q66N97	Q66n97 hepatitis c
45	2636	86.3	571	2		

```
                                   ALIGNMENTS

RESULT 1
Q6B4Q1
ID   Q6B4Q1                 PRELIMINARY;      PRT;    571 AA.
AC   Q6B4Q1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   NS5b (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682462; AAT84167.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR002166; HCV_RdRp.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
FT   NON_TER        1      1
FT   NON_TER      571    571
SQ   SEQUENCE   571 AA;  63223 MW;  975788F1925D0DBC CRC64;

     Query Match            94.0%;  Score 2873;  DB 2;  Length 571;
     Best Local Similarity  95.8%;  Pred. No. 6.9e-218;
     Matches  543;  Conservative  13;  Mismatches  11;  Indels  0;  Gaps  0;

Qy        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
            ||||||||||||||||||||||||||||||||||||||||||||:||||:||||||||||
Db        5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL  64

Qy       65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||||||||:|||:||||||:|||||
Db       65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKATNHIR 124

Qy      125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      125 SVWKDLLEDTDTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184

Qy      185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            |||||||||||||||||||||||||||||||||||||:||||||||||||:|:|||||||
Db      185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVTESDIRVEESIY 244

Qy      245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
```

```
Db            ||||||||||||:|||: :||||||||||||||||||||||||
   245 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 304
Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDEAASLRVFTEAMTRYSAPPGDPPKPEYDL 364
            |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRVFTEAMTRYSAPPGDPPKPEYDL 364
Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||:|||||||:||||||||||||||||||||||||||||||
Db   365 ELITSCSSNVSVAHDAAGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 484
Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 544
Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
            |||:|||||||||||||||||:|| :|
Db   545 AASQLDLSGWFVAGYSGGDIYHSVSHA 571
```

RESULT 2
```
Q66NA8
ID   Q66NA8        PRELIMINARY;        PRT;    571 AA.
AC   Q66NA8;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682754; AAU08311.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER       1      1
FT   NON_TER     571    571
SQ   SEQUENCE   571 AA;    63201 MW;    5590C9F59B69B2A0 CRC64;
```

```
Query Match    93.9%;  Best Local Similarity 95.8%;
Matches 543;  Conservative 14;  Mismatches 10;  Indels 0;  Gaps 0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||:||||||||||||||:|||||||||:| ||||||||
Db     5 SMSYTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASQRQKKVTFDRLQVL 64
Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||||||||||||||||||:||||||||:|||: |:||||||
Db    65 DDHYRDVLKEMKAKASTVKARLLSVEEACKLTPPHSARSKFGYGAKDVRNLSGKAVNHIR 124
Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   125 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
            |||||||||||||||||||||:|||||||| |||||||||||||||||||||||||||||
Db   185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVTESDIRVEESIY 244
```

--- right column ---

```
Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
            ||||||||||||||||||||:||||:||||||||||||||||||||||||||||||:||
Db   245 QCCDLAPEARQAIRSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 304
Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
            ||||||||||||||||:|||||||||||||||:|||||||||||||||||||||:|||||
Db   305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRVFTEAMTRYSAPPGDPPKPEYDL 364
Qy   365 ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
            |||||||||||||||||||:|||||||:||||||||||||||||||||||||||||||
Db   365 ELITSCSSNVSVAHDAAGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
            |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 484
Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
            |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 544
Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
            |||:|||||||||||||||||:||:||
Db   545 AASQLDLSGWFVAGYSGGDIYHSVSHA 571
```

RESULT 3
```
Q66N95
ID   Q66N95        PRELIMINARY;        PRT;    571 AA.
AC   Q66N95;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682768; AAU08324.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER       1      1
FT   NON_TER     571    571
SQ   SEQUENCE   571 AA;    63249 MW;    F76482BE94526FDA CRC64;
```

```
Query Match    93.5%;  Best Local Similarity 94.9%;  Pred. No. 1.3e-216;
Matches 538;  Conservative 17;  Mismatches 12;  Indels 0;  Gaps 0;

Qy     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||||||||||||||||||||||:||||||||||||||:||||||||:|| ||||||||
Db     5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMYATTSRSASQRQKKVTFDRLQVL 64
Qy    65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            |||||||||||||||||||||:||||||||||||||:||||||||:|||::|:|||||||
Db    65 DDHYRDVLKEMKAKASTVKARLLSVEEACKLTPPHSARSKFGYGAKDVRNLSGKAVNHIR 124
Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   125 SVWKDLLEDTDTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
Qy   185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTEDDIRVEESIY 244
            ||||||||||||||||||||||||||||||:|||||||||||||||||||:|||||||||
Db   185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY 244
```

```
Qy   185  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKVPMGFAYDTRCFDSTVTENDIRVEESIY  244
          ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db   185  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKVPMGFAYDTRCFDSTVTENDIRVEESIY  244

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          |||||||||||||||||||||||:|||||||:||||||||||||||||||||||||:|||
Db   245  QCCDLAPEARQAIRSLTERLYIGGPLTNSRGQNCGYRRCRASGVLTTSCGNTLTCYLKAT  304

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
          ||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||||
Db   305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPKPEYDL  364

Qy   365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
          ||||||||||||||::||||||||||||||:||||||||||||||||||||||||||||
Db   365  ELITSCSSNVSVAHDATGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
          ||||:||||||||||||||||||||||||||||||||:|||||:|||||||||||||||
Db   425  ARMVLTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG  484

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
          |||||:|||||||||||||||||||||||||||||||||:|:|||||||||||||||||
Db   485  EINRVAACLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP  544

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
          ||||:|:|||||||||||||||:|:||
Db   545  AASQLDLSGWFVAGYSGGQIYHSVSHA  571
```

RESULT 4

```
Q66NA6
ID   Q66NA6;                PRELIMINARY;   PRT;   571 AA.
AC   Q66NA6;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682756; AAU08313.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER         1        1
FT   NON_TER       571      571
SQ   SEQUENCE   571 AA;   63183 MW;   44SC9E3B0939414 6 CRC64;
```

Query Match   93.5%; Score 2857; DB 2; Pred. No. 1.3e-216; Length 571;
Best Local Similarity   95.2%;
Matches   540; Conservative   15; Mismatches   12; Indels   0; Gaps   0;

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||||||||||||||||||||||||||||||||:||::|:|||:|||||||||||||||
Db     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL   64

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          ||||||||||||||||||||||||||||||||||||||||||||||||||:||:|||||
Db    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSGKAVNHIR  124

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   125  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          |||||||||||||||||  ||| ||||||| || :|||
```

```
Db   185  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKVPMGFAYDTRCFDSTVTENDIRVEESIY  244
          ||||||||||||||||||:|||||||:||:|||:|||||||||||||||||||||||||
Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          |||||||||||||||||||||||||:|||||:|||||||||||||||||||||||||:||
Db   245  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT  304

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
          ||||||||||||||:|||||||||||||||||||||||||||||||||||||||:||||
Db   305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPKPEYDL  364

Qy   365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
          ||||||||||||||::||||||||||||||:|||||||||||||||||||||||||||||
Db   365  ELITSCSSNVSVAHDATGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424

Qy   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
          ||||:||||||||||||||||||||||||||||||||:|||||:||||||||||||||||
Db   425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG  484

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
          |||||:||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   485  EINRVAACLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP  544

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
          ||||:|:||||||||||||||:|:|||
Db   545  AASQLDLSGWFVAGYSGGQIYHSVSHA  571
```

RESULT 5

```
Q66N84
ID   Q66N84                 PRELIMINARY;   PRT;   571 AA.
AC   Q66N84;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682779; AAU08335.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER         1        1
FT   NON_TER       571      571
SQ   SEQUENCE   571 AA;   63393 MW;   CD0FE085358BEE67 CRC64;
```

Query Match   93.5%; Score 2856; DB 2; Length 571;
Best Local Similarity   95.1%; Pred. No. 1.5e-216; Indels   0; Gaps   0;
Matches   539; Conservative   17; Mismatches   11;

```
Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||||||||||||||||||||||||||||||||:||::|:|||:|||||||||||||||
Db     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL   64

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          ||||||||||||||||||||||||||||||||||||||:||||||||||:||:||||||
Db    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSGKATNHIR  124

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db   125  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          ||||||||||||||||| |||||||||
```

```
Db              |||||||||||||||||||| :||||||||                          244
          185 TLPQAVWGSSYGFQYSPAQRVEFLVNAWKSKKCPMGFAYDTRCFDSTVTESDIRVEESIY 244

Qy        QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||||||||||:||||:|||||||||||||||||||||||||||||||||||
Db        QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT 304

Qy        AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364

Qy        ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||:||||||||||||||:||||||||||||||||||||||||||||
Db        ELITSCSSNVSVAHDATGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy        ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||:|||||||||:||||:|||||||||||||||
Db        ARMILMTHFFSILLAQEQLEKALDCQIYGACHSIEPLDLPQIIQRLHGLSAFSLHSYSPG 484

Qy        EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGRYLFNWAVRTKLKLTPIP 544

Qy        AASRLDLSGWFVAGYSGGDIYHSLSRA 571
          |||:|||:||||||||:|||||:
Db        AASQLDLSNWFVAGYSGRDIYHSVSHA 571
```

RESULT 6
Q6B4Q2
ID   Q6B4Q2;
AC   Q6B4Q2;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   NS5b (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682461; AAT84166.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 63123 MW; 931743E7C7819F36 CRC64;

Query Match   93.4%; Score 2853; DB 2; Length 571;
Best Local Similarity   95.2%; Pred. No. 2.6e-216;
Matches 540; Conservative 15; Mismatches 12; Indels 0; Gaps 0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||||:|||||||||:|||:||:|||
Db     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKVGYGAKDVRNLSSKAINHIH 124
```

```
Db              ||||||||||||||||||||                                    184
          125 SVWKDLLEDTDPIQTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184

Qy        SVWKDLLEDTATPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
          ||||||||:|  :||||||||||||||||||||||||||||||||||||||||||||||
Db        TLPQAVWGSSYGFQYSPGQRVEFLVNAWKSKSPMGFAYDTRCFDSTVTENDIRVEESIY 244

Qy        TLPQAVWGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIY 244
          ||||||||||||||||: |||||||: :|||: |:||||||||||||||:|||||||||
Db        QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 304

Qy        QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
          ||||||||||||||||||||:||||:|||||||||||||||||||||||||||||||||:
Db        AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364

Qy        AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
          |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        ELITSCSSNVSVAHDATGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy        ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
          ||||||||||||||||:||||||||||||||:||||||||||||||||||||||||||||
Db        ARMILMTHFFSILLAQEQLEKALDCQIYGACHSIEPLDLPQIERLHGLSAFSLHSYSPG 484

Qy        ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
          |||||||||||||||||||||||||||||:|||||||||:||||:|||||||||||||||
Db        EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGRYLFNWAVRTKLKLTPIP 544

Qy        EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
          ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        AASRLDLSGWFVAGYSGGDIYHSVSHA 571

Qy        AASRLDLSGWFVAGYSGGDIYHSLSRA 571
          ||||||||||||||||:|||||:
Db        AASRLDLSSWFVAGYSGGDIYHSVSHA 571
```

RESULT 7
Q66N79
ID   Q66N79;
AC   Q66N79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682784; AAU08340.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 63317 MW; 4BC8AC1E98153857 CRC64;

Query Match   93.3%; Score 2851; DB 2; Length 571;
Best Local Similarity   95.2%; Pred. No. 3.7e-216;
Matches 540; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

```
Qy      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||||||||||||||||||||||||||||||||:||:|||||||:|||||||||||||
Db      5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL 64

Qy     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |||||||||||||||||||||||||||||||||||||:||||||||||:|||:||||||:
Db     65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKAVNHIH 124
```

Column 1 (RESULT 8)

```
Qy 125  SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
        |||||||  |||  ||||||||||||||||||||||||||||||||||||||||||||||
Db 125  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
Qy 185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
        |||||||||||||||||||  ||:|:|  |||:|  |||||||||||||||||||||||||
Db 185  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKVPMGFAYDTRCFDSTVTENDIRVEESIY  244
Qy 245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
        |||||||||||||||||||||:::||  || |||||||||||||||||||||||||||||
Db 245  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT  304
Qy 305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
        ||||||||||||||||||||||||||||||||||||||||||||||||||| ||: ||||
Db 305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPPKPEYDL 364
Qy 365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
        |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db 365  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
Qy 425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
        |||||||||||||||||:::|||||||||||||||:||||||||||||||||||||||||
Db 425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG  484
Qy 485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
        ||||||||||||||||||||||||||||||||||||:||:||||||||||||||||||||
Db 485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  544
Qy 545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
        ||:||||||||||||||||||:
Db 545  DASRLDLSGWFVAGYSGGHIYHSVSHA  571
```

RESULT 8
Q66N83
ID   Q66N83   PRELIMINARY;   PRT;   571 AA.
AC   Q66N83;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682780; AAU08336.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA;   63177 MW;   8649F9B5ED28DEED CRC64;

```
Query Match             93.2%;  Score 2846;  DB 2;  Length 571;
Best Local Similarity   94.9%;  Pred. No. 9.3e-216;
Matches  538;  Conservative  17;  Mismatches 12;  Indels  0;  Gaps   0;

Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         |||||||||||||||||||||||||||||||||||:||:||||||:||||||||||||||
Db    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL  64
Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSSRAVNHIR 124
         ||| |||
```

Column 2 (RESULT 9)

```
Db   65  DDHYRDVLKEMKAKASTVKAKLLSIEBACKLTPPHSAKSKFGYGAKDVRNLSSKAVKHIR  124
         |||||||||||||||||||||||||| ||| |||||||||||| ||| ||||| |||||
Qy  125  SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   184
         |||||||  |||  ||||||||||||||||  |||||||||||||||||||||||||||
Db  125  SVWEDLLEDTDPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS   184
Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         |||||||||||||||||||  ||:|:|  |||:|  |||||||||||||||||||||||||
Db  185  TLPQAVMGSAYGFQYSPGQRVEFLVNAWKSKRNPMGFSYDTRCFDSTVTESDIRVEESIY  244
Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         ||||||||||||||||||||:::|| || |||||||||||||||||||||||||||||
Db  245  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  304
Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||| ||||||||||||:||||||||||||||||||||||||||||||
Db  365  ELITSCSSNVSVAHDATGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
Qy  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         ||||||||||||||||:::|||||||||||||||:|||||||||:|||||||||||||||
Db  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG  484
Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         |||||||||||||||||||||||||||||||||||:|||:|||||||||||||||||||
Db  485  EINRVASCLRKLGVPPLRARWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 544
Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
         ||:|||||||||||||||||||  ||
Db  545  AASQLDLSGWFVAGYSGGDIYHSGSHA  571
```

RESULT 9
Q66NA2
ID   Q66NA2   PRELIMINARY;   PRT;   571 AA.
AC   Q66NA2;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682760; AAU08317.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA;   63227 MW;   F54BDE550ACDEEBC CRC64;

```
Query Match             92.9%;  Score 2839;  DB 2;  Length 571;
Best Local Similarity   94.5%;  Pred. No. 3.3e-215;
Matches  536;  Conservative  19;  Mismatches 12;  Indels  0;  Gaps   0;

Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         |||||||||||||||||||||||||||||||||||:||:|||||:|||||||||||||||
Db    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMYATTSRSASQRQKKVTFDRLQVL  64
Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSSRAVNHIR 124
         ||| |||
```

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
        ||||||||||||||||||||||||||||:|||||||||| |||:|:||||||||:||||||
Db   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACQUTPPHSARSKFGYGAKDVRSLSSKAVNHIR 124

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125 SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIIFPDLGVRVCEKMALYDVVS 184

Qy 185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
        ||||||||||||||||||||||:||||||:|:||:||||||||||||||||:|||||||||
Db 185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKNPMGFSYDTRCFDSTVTESDIRVEESIY 244

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
        ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||:
Db 245 QCCDLAPEARQAIRSLTERLYIYGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 304

Qy 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
        ||||||||||||||||:|||||||||||||||||||||||||||||||||||||:|||||
Db 305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPKPEYDL 364

Qy 365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
        ||||||||||||:|||||||||||||||||:|||||||||||||||||||||||||||||
Db 365 ELITSCSSNVSVAFDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
        |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 484

Qy 485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
        ||||||||||||||||||||:||||:|||||||||||||||||||||||||||||||||||
Db 485 EINRVASCLRKLGVPPLRVWRHRARNVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP 544

Qy 545 AASRLDLSGWFVAGYSGGGDIYHSLSRA 571
        ||||:||||||||||||||||||||:
Db 545 AASQLDLSGWFVAGYSGGGDIYHSVSHA 571

RESULT 10
Q66N78
ID   Q66N78
AC   Q66N78;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682785; AAU08341.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 63360 MW; 4521656175D8334A CRC64;

Query Match   92.9%; Score 2837; DB 2; Length 571;
Best Local Similarity   94.7%; Pred. No. 4.8e-215;
Matches 537; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
        |||||||||||||||||||||||||||||||||||||||:||||||| ||||||:||||||
Db   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYVTTSRSASQRQKKVTFDRMQVL 64

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
        ||||||||||||||||||||||||||||||||||||||||||:|:||||||||:|||||||
Db   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACQUTPPHSARSKFGYGAKDVRSLSSKAVNHIR 124

Qy 125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
        |||||| :|||:||||||||||||||||||||||||||||||:|||||||||||||||||
Db 125 SVWKDLQEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIIFPDLGVRVCEKMALYDVVS 184

Qy 185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
        ||||||||||||||||||:|||||||||||||:|:||:|||||||||||||:|||||||||
Db 185 TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKNPMGFSYDTRCFDSTVTESDIRVEESIY 244

Qy 245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
        ||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||||:
Db 245 QCCDLAPEARQAIRSLTERLYIYGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 304

Qy 305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
        ||||||||||||||||:|||||||||||||||||||||||||||||||||||||:|||||
Db 305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPKPEYDL 364

Qy 365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
        ||||||||||||:|||||||||||||||||:|||||||||||||||||||||||||||||
Db 365 ELITSCSSNVSVAHDATGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
        |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db 425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPG 484

Qy 485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
        ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db 485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 544

Qy 545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
        |||| |||||||||||||||||:
Db 545 AASQLDLSGWFVAGYSGRHIYHSVSHA 571

RESULT 11
Q66NA0
ID   Q66NA0
AC   Q66NA0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682762; AAU08319.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 63254 MW; 475D83341A64631B CRC64;

Query Match   92.9%; Score 2837; DB 2; Length 571;
Best Local Similarity   94.5%; Pred. No. 4.8e-215;
Matches 536; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
        |||||||||||||||||||||||||||||||||||||||:||||||| ||||||:||||||
Db   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL 64

RESULT 12

```
Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         ||||:|||||||||||||||||||||||||||||||||:||||||||||||:||||:|:|
Db   65  DDHYQDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKAVRHIN  124

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         ||||||||| :||:||||||||||||||||||||||||||||||||||||||||||||||
Db  125  SVWKDLLEDNETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         ||||||||||||||||| :|||||||||:|:|| |||:||||||||||||||||||||||
Db  185  TLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKRNPMGFAYDTRCFDSTVTENDIRVEESIY  244

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         |||||||||||||:||||||||:||:||||||||||||||||||||||||||||||||||:
Db  245  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  304

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||:|||||||||||||||||||:|||||||||||||||||||||||||
Db  305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDPPQPEYDL  364

Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||:|||||| |||||:|||||||||||||||||||||||||||||||
Db  365  ELITSCSSNVSVAHDATGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424

Qy  425  ARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         ||||||||||||||||||||||||||||||||||||||||:|| :|||||||||||||||
Db  425  ARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPLIIQRLHGLSAFSLHSYSPG  484

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
         |||||||||||||||||||||||:|:|
Db  545  AASRLDLSGWFVAGYSGGDIYHSVSHA  571
```

RESULT 12

```
Q66N74
ID   Q66N74
AC   Q66N74;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682789; AAU08345.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER      1      1
FT   NON_TER    571    571
SQ   SEQUENCE   571 AA;  63220 MW;  7461A702D65805AD CRC64;
```

Query Match 92.8%; Score 2836; DB 2; Pred. No. 5.7e-215; Length 571;
Best Local Similarity 94.0%;
Matches 533; Conservative 23; Mismatches 11; Indels 0; Gaps 0;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         ||||||||||||||||||||||||||||||||||||| :|||||||||||||||||||||
```

```
Db    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVL  64

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         ||||:|||||||||||||||||||:||||||||||||||||||||||||:|||:|:|:||
Db   65  DDHYQDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSKAINHIR  124

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         |||:|||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  125  SVWEDLLEDTDTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         |||||||||:|||||||:|||||||||:|:|:| |||:||||||||||||:|||||||||
Db  185  TLPQAVMGSAYGFQYSPGQRVEFLVNAWKSKRTPMGFAYDTRCFDSTVTESDIRVEESIY  244

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||:|:
Db  245  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  304

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||||||||||:|||||||||||||||||||:|||||||||||||||||||||||||
Db  305  AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDPPQPEYDL  364

Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         |||||||||||||||:|||||| |||||:|||||||||||||||||||||||||||||||
Db  365  ELITSCSSNVSVAHDATGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424

Qy  425  ARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         ||||||||||||||||||||||||||||||||||||||||:|| :|||||||||||||||
Db  425  ARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPLIIQRLHGLSAFSLHSYSPG  484

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         |||||||||||||||||||||||||||||||||||||||| ||:||||||||||||||| :
Db  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIS  544

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRA  571
         |||:||||||||||||||| |||:|:|
Db  545  AASQLDLSGWFVAGYSGGQIYHSVSHA  571
```

RESULT 13

```
Q66N89
ID   Q66N89
AC   Q66N89;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682774; AAU08330.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER      1      1
FT   NON_TER    571    571
SQ   SEQUENCE   571 AA;  63492 MW;  111E17000D53EB8C CRC64;
```

Query Match 92.3%; Score 2820; DB 2; Length 571;
Best Local Similarity 94.0%; Pred. No. 1e-213;
Matches 533; Conservative 20; Mismatches 14; Indels 0; Gaps 0;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
         |||||||||||||||||||||||||||||||||||||:|||||||||:||||||||||||
```

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHEHNLVYSTTSRSASLRQKKVTFDRLQVL 64
           ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 64

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           |||||||:|||:||||||||||||||:|||||||||||:|||||||||||||||:|||:|
Db   65 DNHYQDVLKEMKAKASTVKARLLSVEEACQLTPPHSARSKFGYGAKDVRSLSSKAVNHIN 124

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   125 SVWKDLLEDTATPIDTTIMAKNEVFCVDPTKGGRKPARFIVFPDLGVRVCEKMALYDVVS 184

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||||||||:||||||||||:||||||||:|||||||||||||||||||:||||||||
Db   185 TLPQAVMGSAYGFQYSPGQRVDFLVNAWKSKRNPMGFSYDTRCFDSTVTESDIRVEESIY 244

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           |||||||||||||||||||||:|||||||||||||||||||:||||||||||||||:|||
Db   245 QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 304

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db   305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           |||||||||||||||:|||||||||||||::|||||||||||||||||||||||||||||
Db   365 ELITSCSSNVSVAHDATGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy   425 ARMILMTHFFSILLAQELEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   425 ARMILMTHFFSILLAQELDKALDCQIYGACYSIEPLDLPQIQRLHGLSAFSLHSYSPG 484

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP 544

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
           ||||||||||||||||||||||||||
Db   545 AASRLDLSGWFVAGYSGRHIYHSVSHA 571

RESULT 14
Q66NB0
ID   Q66NB0;
AC   Q66NB0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682752; AAU08309.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 63097 MW; C02E7328F3417842 CRC64;

Query Match   91.8%;
Best Local Similarity   93.5%;
Matches 530; Conservative 21; Mismatches 16; Indels 0; Gaps 0;

Qy   5 SMSYTWTGALVTPCAAEESKLPINALSNSLLRHHNVYATTSRSASLRQKKVTFDRMQVL 64
           ||||||||||:||||||||||||||||||||||||||:|:|||||||||||||||||:||
Db   5 SMSYTWTGALVTPCAAEESKLPINALSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVL 64

Qy   65 DDHYRDVLKEMKAKASTVKARLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
           ||||||||||||||||||||||||||||||||||||:|||||||||||||||||:||||:
Db   65 DDHYRDVLKEMKAKASTVKARLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIN 124

Qy   125 SVWKDLLEDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
           ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCERMALYDVVS 184

Qy   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTESDIRVEESIY 244

Qy   245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   245 QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS 304

Qy   305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   305 AACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL 364

Qy   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   365 ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW 424

Qy   425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG 484
           ||||||||||||||||:|||||||||||||||||||||||||||:|||||||||||||||
Db   425 ARMILMTHFFSILLAQEQLEKALDFQIYGACYSIEPLDLPQIQRLHGLSAFSLHSYSPG 484

Qy   485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP 544
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   485 EINRVASSLRKLGVPPLRVWRHRARSVRAKLLSRGGRAATCGKYLFNWAVRTKLRLTPIP 544

Qy   545 AASRLDLSGWFVAGYSGGDIYHSLSRA 571
           |||||||||||||||||||||||||:|
Db   545 AASRLDLSSWFVAGYSGGDIYHSVSHA 571

RESULT 15
Q66NA7
ID   Q66NA7;
AC   Q66NA7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA   Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT   "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT   HCV/HIV Coinfected Subjects.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY682755; AAU08312.1; -.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF00998; Viral_RdRP; 1.
KW   Polyprotein.
FT   NON_TER   1   1
FT   NON_TER   571   571
SQ   SEQUENCE   571 AA; 62829 MW; 1EB09F6368B4B729 CRC64;

Query Match   87.7%;   Score 2679; DB 2; Length 571;
Best Local Similarity   88.7%;   Pred. No. 1.4e-202;
Matches 503; Conservative 29; Mismatches 35; Indels 0; Gaps 0;

```
Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL   64
          ||||:||||||:||||||:||||||||||||||||||||||||||||| :|||||||||||
Db    5 SMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACCQRQKKVTFDRLQVL   64

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
         |:|||||||:|||:|:|||:|||||||||:|||||||||||||||||||||  : :||:| 
Db   65 DSHYRDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRSHARKAVAHIN  124

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
         |||||||||:.||:||||||||||||||||||||||||||||||||||||||||||||||
Db  125 SVWKDLLEDSVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184

Qy  185 TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
         .||||||||||||||||.|||||||:..||:||.|||||||||||||||||.|||.|.|||
Db  185 KLPQAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIY  244

Qy  245 QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
         |||||.|:||.||:|||||||||||:||:.|:|||||||||||||||||||||||||:||.
Db  245 QCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKAQ  304

Qy  305 AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
         ||||||.||||||||.|||||||||||||.|||||.|||||||||||||||||||||||||
Db  305 AACRAAGLQDCTMLVCGDDLVVICESAGVQEDAAASLRAFTEAMTRYSAPPGDPPQPEYDL  364

Qy  365 ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
         ||||||||||||||||.|||||||||||||.||||||||||||||||||||||||||.||||
Db  365 ELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLW  424

Qy  425 ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
         |||||||||||:||||:.||:|||||:|||||||||||||||.||:|||||||||||||||
Db  425 ARMILMTHFFSVLIARDQLEQALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPG  484

Qy  485 EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
         |||||:.|||||||||||.|||||||||:|||.|||||||||||||||||||||||||||.
Db  485 EINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFNWAVRTKLKLTPIA  544

Qy  545 AASRLDLSGWFVAGYSGGDIYHSLSRA  571
         ||.:||||||||:||||||||||.|.|
Db  545 AAGQLDLSGWFTAGYSGGDIYHSVSHA  571
```

Search completed: September 22, 2005, 14:54:01
Job time : 181 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:      September 22, 2005, 14:48:15 ; Search time 43 Seconds
                                  (without alignments)
                             1006.894 Million cell updates/sec

Title:           US-10-712-479-4
Perfect score:   3055
Sequence:     1 MASMSMSYTWTGALITPCAA.........GGDIYHSLSRARPRHHHHH 580

Scoring table:   BLOSUM62
             Gapop 10.0 , Gapext 0.5

Searched:    513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:     489056

Minimum DB seq length: 0
Maximum DB seq length: 585

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Issued Patents AA:*
             1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
             2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
             3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
             4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
             5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
             6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                             SUMMARIES
```

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2879	94.2	576	4	US-09-541-990A-1	Sequence 1, Appli
2	1124	36.8	496	3	US-08-867-611-10	Sequence 10, Appl
3	1124	36.8	496	4	US-09-690-359-10	Sequence 10, Appl
4	1124	36.8	496	5	PCT-US92-06965A-15	Sequence 15, Appl
5	908	29.7	504	1	US-07-853-985A-6	Sequence 6, Appli
6	908	29.7	504	1	US-07-681-703B-6	Sequence 6, Appli
7	908	29.7	504	2	US-08-184-236-6	Sequence 6, Appli
8	908	29.7	504	2	US-08-407-410B-6	Sequence 6, Appli
9	908	29.7	504	2	US-08-485-500-6	Sequence 6, Appli
10	908	29.7	504	5	PCT-US91-02370-6	Sequence 6, Appli
11	908	29.7	504	5	PCT-US94-04174-6	Sequence 6, Appli
12	884	28.9	187	1	US-07-853-985A-2	Sequence 2, Appli
13	884	28.9	187	1	US-07-681-703B-2	Sequence 2, Appli
14	884	28.9	187	1	US-08-184-236-2	Sequence 2, Appli
15	884	28.9	187	2	US-08-407-410B-2	Sequence 2, Appli
16	884	28.9	187	2	US-08-485-500-2	Sequence 2, Appli
17	884	28.9	187	5	PCT-US91-02370-2	Sequence 2, Appli
18	884	28.9	187	5	PCT-US94-04174-2	Sequence 2, Appli
19	759	24.8	152	3	US-08-444-818-62	Sequence 62, Appl
20	748	24.5	171	3	US-08-444-818-58	Sequence 58, Appl
21	686	22.5	387	3	US-08-867-611-12	Sequence 12, Appl
22	686	22.5	387	4	US-09-690-359-12	Sequence 12, Appl
23	686	22.5	387	5	PCT-US92-06965A-17	Sequence 17, Appl
24	678	22.2	156	5	PCT-US94-04174-19	Sequence 19, Appl
25	621	20.3	133	3	US-08-444-818-60	Sequence 60, Appl
26	613.5	20.1	393	3	US-08-867-611-14	Sequence 14, Appl
27	613.5	20.1	393	4	US-09-690-359-14	Sequence 14, Appl
28	613.5	20.1	393	5	PCT-US92-06965A-19	Sequence 19, Appl
29	584	19.1	516	1	US-08-638-911A-41	Sequence 41, Appl
30	570	18.7	113	3	US-08-256-568B-82	Sequence 82, Appl
31	570	18.7	113	3	US-09-038-369B-82	Sequence 82, Appl
32	570	18.7	113	4	US-09-378-900A-82	Sequence 82, Appl
33	570	18.7	113	4	US-09-899-044-82	Sequence 82, Appl
34	570	18.7	113	4	US-09-878-281A-214	Sequence 214, App
35	549	18.0	113	3	US-08-836-075A-58	Sequence 58, Appl
36	536	17.5	113	3	US-08-836-075A-54	Sequence 54, Appl
37	535	17.5	113	3	US-08-836-075A-56	Sequence 56, Appl
38	531	17.4	113	3	US-08-836-075A-62	Sequence 62, Appl
39	521	17.1	113	3	US-08-836-075A-68	Sequence 68, Appl
40	515	16.9	267	3	US-08-454-928-10	Sequence 10, Appl
41	499	16.3	113	4	US-08-836-075A-64	Sequence 64, Appl
42	467	15.3	100	4	US-08-469-260A-396	Sequence 396, App
43	467	15.3	100	4	US-08-488-446-396	Sequence 396, App
44	467	15.3	100	4	US-08-467-344A-396	Sequence 396, App
45	467	15.3	100	4	US-08-424-550B-396	Sequence 396, App

```
                         ALIGNMENTS

RESULT 1
US-09-541-990A-1
; Sequence 1, Application US/09541990A
; Patent No. 6433489
; GENERAL INFORMATION:
;  APPLICANT:   Lesburg, Charles A.
;  APPLICANT:   Cable, Michael
;  APPLICANT:   Hong, Zhi
;  APPLICANT:   Mannarino, Anthony
;  APPLICANT:   Weber, Patricia
;  TITLE OF INVENTION:  Compositions of Hepatitis C Virus NS5B Polymerase and
;  TITLE OF INVENTION:  Methods for Crystallizing Same
;  FILE REFERENCE:  IN0988 US
;  CURRENT APPLICATION NUMBER: US/09/541,990A
;  CURRENT FILING DATE:  2000-04-03
;  NUMBER OF SEQ ID NOS:  1
;  SOFTWARE:   PatentIn Ver. 2.1
; SEQ ID NO 1
;  LENGTH: 576
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
;  FEATURE:
;  NAME/KEY: VARIANT
;  LOCATION: (335)
;  OTHER INFORMATION:  amino acids at 335, 344 and 550 may be threonine
;  OTHER INFORMATION:  or valine at 335, valine or alanine at 344 and
;  OTHER INFORMATION:  arginine or glutamine at 550
US-09-541-990A-1

  Query Match         94.2%;  Score 2879;  DB 4;  Length 576;
  Best Local Similarity   96.0%;  Pred. No. 1.2e-286;
  Matches  545;  Conservative  11;  Mismatches  12;  Indels  0;  Gaps  0;

Qy      7  SYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVLDD  66
           |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db      9  SYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDD  68

Qy     67  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSV  126
           |||||||||||||||||||||||||||||||||||||||||:||:|||||  ||
Db     69  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIHSV  128

Qy    127  WKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  186
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129  WKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  188

Qy    187  PQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCCPMGFSYDTRCFDSTVTENDIRVEESIYQC  246
           |||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db    189  PQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIYQC  248

Qy    247  CDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAA  306
```

```
Db  249 CDLAPEARQAIKSLTERLYIIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAA 308
Qy  307 CRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLEL  366
Db  309 CRAAKLQDCTMLVNGDDLVVICESAGXQEDAASLRXFTEAMTRYSAPPGDPPQPEYDLEL  368
Qy  367 ITSCSSNVSVAHDASGKRVVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  426
Db  369 ITSCSSNVSVAHDASGKRVVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  428
Qy  427 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI  486
Db  429 MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI  488
Qy  487 NRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAA  546
Db  489 NRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAA  548
Qy  547 SRLDLSGWFVAGYSGGDIYHSLSRARPR 574
Db  549 SXLDLSGWFVAGYSGGDIYHSLSRARPR 576
```

RESULT 2
US-08-867-611-10
; Sequence 10, Application US/08867611
; Patent No. 6172189
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
TITLE OF INVENTION: ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565

FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-867-611-10

```
Query Match             36.8%;  Score 1124;  DB 3;  Length 496;
Best Local Similarity   84.8%;  Pred. No. 2.4e-106;
Matches  212;  Conservative  14;  Mismatches  24;  Indels   0;  Gaps   0;

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKPGYGAKDVRSLLSSRAVNHI 123
Db  247 LDSHYQDVLKEVKAASKVKANLLSVEEACSLTPPHSAKSKPGYGAKDVRCHARKAVTHI   306
Qy  124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV  183
Db  307 NSVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV  366
Qy  184 STLPQAVMGSSYGFQYSPKQRVBFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVBESI  243
Db  367 TKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEAI   426
Qy  244 YQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA  303
Db  427 YQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKA  486
Qy  304 AAACRAAKLQ 313
Db  487 RAACRAAGLQ 496
```

RESULT 3
US-09-690-359-10
; Sequence 10, Application US/09690359
; Patent No. 6593083
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
          DESAI, SURESH M
          CASEY, JAMES M
          DAILEY, STEPHEN H
          DAWSON, GEORGE J
          GUTIERREZ, ROBIN A
          LESNIEWSKI, RICHARD R
          STEWART, JAMES L
          RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
                    ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/09/690,359
   FILING DATE: 17-Oct-2000
   CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
   APPLICATION NUMBER: 08/867,611
   FILING DATE: 02-JUN-1997
   APPLICATION NUMBER: US/08/646,757
   FILING DATE: <Unknown>
   APPLICATION NUMBER: US/08/179,896
   FILING DATE: <Unknown>
   APPLICATION NUMBER: US 07/572,822
   FILING DATE: 24-AUG-1990
   APPLICATION NUMBER: US 07/614,069
   FILING DATE: 07-NOV-1990
   APPLICATION NUMBER: US 07/748,561
   FILING DATE: 21-AUG-1991
   APPLICATION NUMBER: US 07/748,565
   FILING DATE: 21-AUG-1991
   APPLICATION NUMBER: US 07/748,566
   FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
   NAME: POREMBSKI, PRISCILLA E
   REGISTRATION NUMBER: 33,207
   REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
   TELEPHONE: 708-937-9556
   TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 10:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 496 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
   MOLECULE TYPE: protein
   SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-690-359-10

Query Match   36.8%; Score 1124; DB 4; Length 496;
Best Local Similarity   84.8%; Pred. No. 2.4e-106;
Matches 212; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHI 123
Db   247 LDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHI 306

Qy   124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 183
Db   307 NSVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 366

Qy   184 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
Db   367 TKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAI 426

Qy   244 YQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 303
Db   427 YQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKA 486

Qy   304 AAACRAAKLQ 313
Db   487 RAACRAAGLQ 496

RESULT 4
PCT-US92-06965A-15
; Sequence 15, Application PC/TUS9206965A
; GENERAL INFORMATION:
;   APPLICANT: DEVARE, S.
;   APPLICANT: DESAI, S.
;   APPLICANT: DAILEY, S.
;   TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ABBOTT LABORATORIES
      STREET: ONE ABBOTT PARK ROAD
      CITY: ABBOTT PARK
      STATE: ILLINOIS
      COUNTRY: U.S.
      ZIP: 60065-3500
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US92/06965A
      FILING DATE: 19920821
      CLASSIFICATION:
   ATTORNEY/AGENT INFORMATION:
      NAME: POREMBSKI, PRISCILLA E.
      REGISTRATION NUMBER: 33,207
      REFERENCE/DOCKET NUMBER: 4834PC.02
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 708-937-6365
      TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 496 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
PCT-US92-06965A-15

Query Match   36.8%; Score 1124; DB 5; Length 496;
Best Local Similarity   84.8%; Pred. No. 2.4e-106;
Matches 212; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

Qy   64 LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHI 123
Db   247 LDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHI 306

Qy   124 RSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 183
Db   307 NSVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVV 366

Qy   184 STLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESI 243
Db   367 TKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAI 426

Qy   244 YQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKA 303
Db   427 YQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKA 486

Qy   304 AAACRAAKLQ 313
Db   487 RAACRAAGLQ 496

RESULT 5
US-07-853-985A-6
; Sequence 6, Application US/07853985A
; Patent No. 5436318
; GENERAL INFORMATION:
;   APPLICANT: Reyes, Gregory
;   APPLICANT: Kim, Jungsuh P.
;   APPLICANT: Moeckli, Randolph
;   TITLE OF INVENTION: Hepatitis C Virus Epitopes
;   NUMBER OF SEQUENCES: 10
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Peter J. Dehlinger
;   STREET: P.O. BOX 60850
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94306
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,985A
FILING DATE: 19920320
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,611
FILING DATE: 06-APR-1990
APPLICATION NUMBER: US 07/594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Dehlinger, Peter J.
REGISTRATION NUMBER: 28,006
REFERENCE/DOCKET NUMBER: 4600-0076.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-324-0880
TELEFAX: 415-324-0960
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-681-703B-6

Query Match   29.7%;   Score 908; DB 1; Length 504;
Best Local Similarity 87.2%;   Pred. No. 3.9e-84;
Matches 177; Conservative   7; Mismatches   19; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAABESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
             ||| | || |||| || | ||||||||||||||||||||||| :||: || ||||||||
Db   302 SMSYSWTGALVTPCAABEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR 124
             ||||::||||| |:|| | | |||||||||||||||||||||| |||| ::|| : ||:
Db   362 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||: : ||
Db   422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 481

Qy   185 TLPQAVMGSSYGFQYSPKQRVBEF 207
             |||||||||||||||||||
Db   482 KLPLAVMGSSYGFQYSPGQRVBEF 504

RESULT 6
US-07-681-703B-6
; Sequence 6, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/681,703B
FILING DATE: 05-APR-1991

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 505,611
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 594,854
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-076.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-681-703B-6

Query Match   29.7%;   Score 908; DB 1; Length 504;
Best Local Similarity 87.2%;   Pred. No. 3.9e-84;
Matches 177; Conservative   7; Mismatches   19; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAABESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
             ||| | || |||| || | ||||||||||||||||||||||| :||: || ||||||||
Db   302 SMSYSWTGALVTPCAABEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIR 124
             ||||::||||| |:|| | | |||||||||||||||||||||| |||| ::|| : ||:
Db   362 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy   125 SVWKDLLEDTDPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||: : ||
Db   422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 481

Qy   185 TLPQAVMGSSYGFQYSPKQRVBEF 207
             |||||||||||||||||||
Db   482 KLPLAVMGSSYGFQYSPGQRVBEF 504

RESULT 7
US-08-184-236-6
; Sequence 6, Application US/08184236
; Patent No. 5538865
; GENERAL INFORMATION:
APPLICANT: Reyes, Gregory
APPLICANT: Kim, Jungsuh P.
APPLICANT: Moeckli, Randolph
TITLE OF INVENTION: Hepatitis C Virus Epitopes
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSE: Dehlinger & Associates
STREET: P.O. BOX 60850
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,236
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/505,611
FILING DATE: 06-APR-1990
APPLICATION NUMBER: US 07/594,854
FILING DATE: 09-OCT-1990

```
        APPLICATION NUMBER: US 07/853,985
        FILING DATE: 20-MAR-1992
        ATTORNEY/AGENT INFORMATION:
        NAME: Fabian, Gary R.
        REGISTRATION NUMBER: 33,875
        REFERENCE/DOCKET NUMBER: 4600-0113
        TELECOMMUNICATION INFORMATION:
        TELEPHONE: 415-324-0880
        TELEFAX: 415-324-0960
INFORMATION FOR SEQ ID NO: 6:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 504 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-184-236-6

Query Match       29.7%;  Score 908; DB 1; Length 504;
Best Local Similarity  87.2%;  Pred. No. 3.9e-84;
Matches 177;  Conservative  7;  Mismatches 19;  Indels 0;  Gaps 0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||:|||:||||||||||||||||||||||||||||   ||||||||||||||
Db  302 SMSYSWTGALVTPCAAEBQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |:||||||:|||||| :|||:||||||| :||||||||||||||||||||  ::||:|
Db  362 DSHYQDVLKEVKAAASKVKANLLSVEBACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVS 184
          |||||||| ||||:||||||||||||||||||||||||||||||||||||||||||:
Db  422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVT 481

Qy  185 TLPQAVMGSSYGFQYSPKQRVEF 207
          :||:|||||||||||||| ||||
Db  482 KLPLAVMGSSYGFQYSPGQRVEF 504


RESULT 8
US-08-407-410B-6
; Sequence 6, Application US/08407410B
; Patent No. 5843636
; GENERAL INFORMATION:
        APPLICANT: Reyes, Gregory
        APPLICANT: Kim, Jungsuh P.
        APPLICANT: Moeckli, Randolph
        APPLICANT: Simonsen, Christian C.
        TITLE OF INVENTION: Hepatitis C Virus Epitopes
        NUMBER OF SEQUENCES: 26
        CORRESPONDENCE ADDRESS:
        ADDRESSEE: Peter J. Dehlinger
        STREET: 350 Cambridge Ave., Suite 100
        CITY: Palo Alto
        STATE: CA
        COUNTRY: USA
        ZIP: 94306
        COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
        CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/407,410B
        FILING DATE:
        CLASSIFICATION: 435
        PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 505,611
        FILING DATE: 06-APR-1990
        PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 594,854
        FILING DATE: 09-OCT-1990
        ATTORNEY/AGENT INFORMATION:
        NAME: Fabian, Gary R.
```

```
        REGISTRATION NUMBER: 33,875
        REFERENCE/DOCKET NUMBER: 4600-076.21
        TELECOMMUNICATION INFORMATION:
        TELEPHONE: (415) 323-8302
INFORMATION FOR SEQ ID NO: 6:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 504 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
US-08-407-410B-6

Query Match       29.7%;  Score 908; DB 2; Length 504;
Best Local Similarity  87.2%;  Pred. No. 3.9e-84;
Matches 177;  Conservative  7;  Mismatches 19;  Indels 0;  Gaps 0;

Qy    5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
          |||||:|||:||||||||||||||||||||||||||||   ||||||||||||||
Db  302 SMSYSWTGALVTPCAAEBQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
          |:||||||:|||||| :|||:||||||| :||||||||||||||||||||  ::||:|
Db  362 DSHYQDVLKEVKAAASKVKANLLSVEBACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVS 184
          |||||||| ||||:||||||||||||||||||||||||||||||||||||||||||:
Db  422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVT 481

Qy  185 TLPQAVMGSSYGFQYSPKQRVEF 207
          :||:|||||||||||||| ||||
Db  482 KLPLAVMGSSYGFQYSPGQRVEF 504


RESULT 9
US-08-485-500-6
; Sequence 6, Application US/08485500
; Patent No. 5843639
; GENERAL INFORMATION:
        APPLICANT: Reyes, Gregory
        APPLICANT: Kim, Jungsuh P.
        APPLICANT: Moeckli, Randolph
        APPLICANT: Simonsen, Christian C.
        TITLE OF INVENTION: Hepatitis C Virus Epitopes
        NUMBER OF SEQUENCES: 26
        CORRESPONDENCE ADDRESS:
        ADDRESSEE: Peter J. Dehlinger
        STREET: 350 Cambridge Ave., Suite 100
        CITY: Palo Alto
        STATE: CA
        COUNTRY: USA
        ZIP: 94306
        COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
        CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/485,500
        FILING DATE:
        CLASSIFICATION: 435
        PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US/08/407,410
        FILING DATE:
        APPLICATION NUMBER: US 505,611
        FILING DATE: 06-APR-1990
        PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 594,854
        FILING DATE: 09-OCT-1990
        ATTORNEY/AGENT INFORMATION:
        NAME: Fabian, Gary R.
        REGISTRATION NUMBER: 33,875
        REFERENCE/DOCKET NUMBER: 4600-076.21
        TELECOMMUNICATION INFORMATION:
```

; MOLECULE TYPE: protein
PCT-US91-02370-6

Query Match   29.7%;   Score 908; DB 5; Length 504;
Best Local Similarity 87.2%;   Pred. No. 3.9e-84;
Matches 177; Conservative   7; Mismatches   19; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||:||:|||||||||:||||||||||||||||||||||||||||||||||| ||||||||||
Db   302 SMSYSWTGALVTPCAABEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            :|||||||||||||||||||||| |||||||||||||||||||||||||:|||::|| : |:|
Db   362 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||:
Db   422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 481

Qy   185 TLPQAVMGSSYGFQYSPKQRVEF 207
            |||||||||||||||||:||||||
Db   482 KLPLAVMGSSYGFQYSPGQRVEF 504

RESULT 11
PCT-US94-04174-6
; Sequence 6, Application PC/TUS9404174
; GENERAL INFORMATION:
;   APPLICANT: Reyes, Gregory
;   APPLICANT: Kim, Jungsuh P.
;   APPLICANT: Moeckli, Randolph
;   TITLE OF INVENTION: HCV Immunodiagnostic Antigens and Antibodies
;   NUMBER OF SEQUENCES: 28
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Law Offices of Peter J. Dehlinger
;   STREET: 350 Cambridge Avenue, Suite 300
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94306
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.24
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US94/04174
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 505,611
;   FILING DATE: 06-APR-1990
;   TELEFAX: 415-324-0960
;   INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 504 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
PCT-US94-04174-6

Query Match   29.7%;   Score 908; DB 5; Length 504;
Best Local Similarity 87.2%;   Pred. No. 3.9e-84;
Matches 177; Conservative   7; Mismatches   19; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||:||:|||||||||:|||||||||||||||||||||||||||||||||||| ||||||||||

;   TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 504 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-485-500-6

Query Match   29.7%;   Score 908; DB 2; Length 504;
Best Local Similarity 87.2%;   Pred. No. 3.9e-84;
Matches 177; Conservative   7; Mismatches   19; Indels   0; Gaps   0;

Qy   5 SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL 64
            |||:||:|||||||||:||||||||||||||||||||||||||||||||||| ||||||||||
Db   302 SMSYSWTGALVTPCAABEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361

Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
            :|||||||||||||||||||||| |||||||||||||||||||||||||:|||::|| : |:|
Db   362 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421

Qy   125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
            |||||||||||:|||||||||||||||||||||||||||||||||||||||||||:
Db   422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 481

Qy   185 TLPQAVMGSSYGFQYSPKQRVEF 207
            |||||||||||||||||:||||||
Db   482 KLPLAVMGSSYGFQYSPGQRVEF 504

RESULT 10
PCT-US91-02370-6
; Sequence 6, Application PC/TUS9102370
; GENERAL INFORMATION:
;   APPLICANT: Reyes, Gregory
;   APPLICANT: Kim, Jungsuh P.
;   APPLICANT: Moeckli, Randolph
;   APPLICANT: Simonsen, Christian C.
;   TITLE OF INVENTION: Hepatitis C Virus Epitopes
;   NUMBER OF SEQUENCES: 26
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Peter J. Dehlinger
;   STREET: P.O. BOX 60850
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US91/02370
;   FILING DATE: 19910405
;   CLASSIFICATION: 435.5
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 505,611
;   FILING DATE: 06-APR-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 594,854
;   FILING DATE: 09-OCT-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fabian, Gary R.
;   REGISTRATION NUMBER: 33,875
;   REFERENCE/DOCKET NUMBER: 4600-076.41
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 504 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear

```
Db  302 SMSYSWTGALVTPCAAEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL 361
        ||   | |:| | |: ::  |:|  |||:|||::||||||| :|||:||||||||:| |
Qy   65 DDHYRDVLKEMKAKASTVKAKLLSVEBACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
        : :|: ||||:| ||| ||||:|||||||:| |||||||||||||||| :|  ||  ||
Db  362 DSHYQDVLKEVKAAASKVKANLLSVEBACSLTPPHSAKSKFGYGAKDVRCHARKAVTHIN 421
Qy  125 SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYDVVS 184
        |||||||||:.||| |||||||||||||||||||||||||||||||||||||||||||
Db  422 SVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVT 481
Qy  185 TLPQAVMGSSYGFQYSPKQRVBF 207
        .|| ||||||||||||||||||
Db  482 KLPLAVMGSSYGFQYSPGQRVBF 504
```

RESULT 12
US-07-853-985A-2
```
; Sequence 2, Application US/07853985A
; Patent No. 5436318
; GENERAL INFORMATION:
;  APPLICANT: Reyes, Gregory
;  APPLICANT: Kim, Jungsuh P.
;  APPLICANT: Moeckli, Randolph
;  TITLE OF INVENTION: Hepatitis C Virus Epitopes
;  NUMBER OF SEQUENCES: 10
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Peter J. Dehlinger
;   STREET: P.O. BOX 60850
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: US
;   ZIP: 94306
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.24
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/853,985A
;   FILING DATE: 19920320
;   CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/505,611
;   FILING DATE: 06-APR-1990
;   APPLICATION NUMBER: US 07/594,854
;   FILING DATE: 09-OCT-1990
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Dehlinger, Peter J.
;   REGISTRATION NUMBER: 28,006
;   REFERENCE/DOCKET NUMBER: 4600-0076.22
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-324-0880
;   TELEFAX: 415-324-0960
; INFORMATION FOR SEQ ID NO: 2:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 187 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-853-985A-2

Query Match             28.9%;   Score 884;   DB 1;   Length 187;
Best Local Similarity   87.6%;   Pred. No. 2.2e-82;
Matches  163;  Conservative  11;  Mismatches  12;  Indels  0;  Gaps  0;

Qy  206 EFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
        |||| .||:||.||||||||||||||||:||.|.||:||||||| |:|| ||:||||||
Db    1 EFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLY 60
Qy  266 VGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYIKARAAACRAAKLQDCTMLVNGDDLV 325
        ||||:||:.||||||||||||||||||||||||||||||| |:|||.|||||||| |||||
Db   61 VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLV 120
```

```
Db  326 VICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRV 385
        ||||||.|||||||||.|||||||||||||||||||||||||||||||||||||.:|||
Qy  121 VICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDGAGKRV 180
Qy  386 YYLTRD 391
        |||||:
Db  181 YYLTRE 186
```

RESULT 13
US-07-681-703B-2
```
; Sequence 2, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
;  APPLICANT: Reyes, Gregory
;  APPLICANT: Kim, Jungsuh P.
;  APPLICANT: Moeckli, Randolph
;  TITLE OF INVENTION: Hepatitis C Virus Epitopes
;  NUMBER OF SEQUENCES: 55
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dehlinger & Associates
;   STREET: 350 Cambridge Ave., Suite 250
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0,  Version #1.25
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/681,703B
;   FILING DATE: 05-APR-1991
;   CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 505,611
;   FILING DATE: 06-APR-1990
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 594,854
;   FILING DATE: 09-OCT-1990
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Fabian, Gary R.
;   REGISTRATION NUMBER: 33,875
;   REFERENCE/DOCKET NUMBER: 4600-076.21
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 324-0880
; INFORMATION FOR SEQ ID NO: 2:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 187 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-681-703B-2

Query Match             28.9%;   Score 884;   DB 1;   Length 187;
Best Local Similarity   87.6%;   Pred. No. 2.2e-82;
Matches  163;  Conservative  11;  Mismatches  12;  Indels  0;  Gaps  0;

Qy  206 EFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
        |||| .||:||.||||||||||||||||:||.|.||:||||||| |:|| ||:||||||
Db    1 EFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLY 60
Qy  266 VGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYIKARAAACRAAKLQDCTMLVNGDDLV 325
        ||||:||:.||||||||||||||||||||||||||||||| |:|||.|||||||| |||||
Db   61 VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLV 120
Qy  326 VICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRV 385
        ||||||.|||||||||.|||||||||||||||||||||||||||||||||||||.:|||
Db  121 VICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDGAGKRV 180
Qy  386 YYLTRD 391
        |||||:
```

Db   181 YYLTRE 186

RESULT 14
US-08-184-236-2
; Sequence 2, Application US/0184236
; Patent No. 5538865
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsuh P.
; APPLICANT: Moeckli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. BOX 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,236
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/505,611
; FILING DATE: 06-APR-1990
; APPLICATION NUMBER: US 07/594,854
; FILING DATE: 09-OCT-1990
; APPLICATION NUMBER: US 07/853,985
; FILING DATE: 20-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-324-0880
; TELEFAX: 415-324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;US-08-184-236-2

Query Match   28.9%; Score 884; DB 1; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```
Qy 206 EFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
          ||:|| ||||||||||||||||||||||:|||||||:|| :||||| ||:||||||||
Db   1 EFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTERAIYQCCDLDPQARVAIKSLTERLY 60

Qy 266 VGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLV 325
          |||| |||:||:|||||||||||||||||||||||:|||| ||||||||||||||:|||
Db  61 VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLV 120

Qy 326 VICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRV 385
          ||||||:|||||||| ||:||||||||||||||||||||||||||||||||||:  |||
Db 121 VICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDGAGKRV 180

Qy 386 YYLTRD 391
          ||||:
Db 181 YYLTRE 186
```

RESULT 15

US-08-407-410B-2
; Sequence 2, Application US/08407410B
; Patent No. 5843636
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsuh P.
; APPLICANT: Moeckli, Randolph
; APPLICANT: Simonsen, Christian C.
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Peter J. Dehlinger
; STREET: 350 Cambridge Ave., Suite 100
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,410B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 323-8302
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-407-410B-2

Query Match   28.9%; Score 884; DB 2; Length 187;
Best Local Similarity 87.6%; Pred. No. 2.2e-82;
Matches 163; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```
Qy 206 EFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLY 265
          ||:|| ||||||||||||||||||||||:|||||||:|| :||||| ||:||||||||
Db   1 EFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTERAIYQCCDLDPQARVAIKSLTERLY 60

Qy 266 VGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLV 325
          |||| |||:||:|||||||||||||||||||||||:|||| ||||||||||||||:|||
Db  61 VGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLV 120

Qy 326 VICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRV 385
          ||||||:|||||||| ||:||||||||||||||||||||||||||||||||||:  |||
Db 121 VICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDGAGKRV 180

Qy 386 YYLTRD 391
          ||||:
Db 181 YYLTRE 186
```

Search completed: September 22, 2005, 14:55:37
Job time : 44 secs

OM protein - protein search, using sw model

    GenCore version 5.1.6
    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: September 22, 2005, 14:49:36 ; Search time 173 Seconds
                 (without alignments)
                 1364.549 Million cell updates/sec

Title: US-10-712-479-4
Perfect score: 3055
Sequence: 1 MASMSMSYTWTGALITPCAA..........GGDIYHSLSRARPRHHHHH 580

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 1826521 seqs, 407012169 residues

Total number of hits satisfying chosen parameters: 1677200

Minimum DB seq length: 0
Maximum DB seq length: 585

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications AA:*
    1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
    2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
    3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
    4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
    5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
    6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
    7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
    8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
    9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
   10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
   11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
   12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
   13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
   14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
   15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
   16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
   17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
   18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
   19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
   20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
   21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
   22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	3055	100.0	580	16	US-10-712-479-4	Sequence 4, Appli
2	2958	96.8	576	17	US-10-842-046-1	Sequence 1, Appli
3	2936	96.1	578	14	US-10-211-455-1	Sequence 1, Appli
4	2936	96.1	578	16	US-10-471-164-3	Sequence 3, Appli
5	2879	94.2	576	14	US-10-471-164-1	Sequence 1, Appli
6	2734.5	89.5	545	16	US-10-471-164-4	Sequence 4, Appli
7	1391	45.5	300	9	US-09-921-397-107	Sequence 17, Appl
8	1283	42.0	352	17	US-09-927-520-17	Sequence 109, App
9	1232	40.3	260	9	US-09-921-397-109	Sequence 3, Appli
10	975.5	31.9	576	14	US-10-211-455-3	Sequence 108, App
11	915	30.0	199	9	US-09-921-397-108	Sequence 108, App
12	623	20.4	127	9	US-09-921-397-110	Sequence 110, App
13	585	19.1	137	9	US-09-921-397-106	Sequence 106, App
14	570	18.7	113	9	US-09-294-121A-82	Sequence 82, Appl
15	570	18.7	113	9	US-09-899-082A-82	Sequence 82, Appl
16	570	18.7	113	9	US-09-899-302-82	Sequence 82, Appl
17	570	18.7	113	10	US-09-899-046-214	Sequence 214, App
18	570	18.7	113	10	US-09-878-281-214	Sequence 214, App
19	570	18.7	113	10	US-09-873-224-214	Sequence 214, App
20	570	18.7	113	10	US-09-899-044-82	Sequence 82, Appl
21	570	18.7	113	16	US-10-822-711-82	Sequence 82, Appl
22	549	18.0	113	9	US-09-851-138-58	Sequence 58, Appl
23	548	17.9	120	14	US-10-292-129-17	Sequence 17, Appl
24	536	17.5	113	9	US-09-851-138-54	Sequence 54, Appl
25	535	17.5	113	9	US-09-851-138-56	Sequence 56, Appl
26	531	17.4	113	9	US-09-851-138-62	Sequence 62, Appl
27	521	17.1	113	9	US-09-851-138-68	Sequence 68, Appl
28	499	16.3	113	9	US-09-851-138-64	Sequence 64, Appl
29	492	16.1	102	9	US-09-921-397-31	Sequence 31, Appl
30	467	15.3	100	8	US-08-424-550B-396	Sequence 396, App
31	464	15.2	113	9	US-09-294-121A-92	Sequence 92, Appl
32	464	15.2	113	9	US-09-899-082A-92	Sequence 92, Appl
33	464	15.2	113	9	US-09-899-302-92	Sequence 92, Appl
34	464	15.2	113	10	US-09-899-046-160	Sequence 160, App
35	464	15.2	113	10	US-09-878-281-160	Sequence 160, App
36	464	15.2	113	10	US-09-899-044-92	Sequence 92, Appl
37	464	15.2	113	10	US-09-873-224-160	Sequence 160, App
38	464	15.2	113	16	US-10-822-711-92	Sequence 92, Appl
39	445	14.6	113	9	US-09-294-121A-89	Sequence 89, Appl
40	445	14.6	113	9	US-09-899-082A-89	Sequence 89, Appl
41	445	14.6	113	9	US-09-899-302-89	Sequence 89, Appl
42	445	14.6	113	10	US-09-899-046-113	Sequence 113, App
43	445	14.6	113	10	US-09-878-281-113	Sequence 113, App
44	445	14.6	113	10	US-09-899-044-89	Sequence 89, Appl
45	445	14.6	113	10	US-09-873-224-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-10-712-479-4
; Sequence 4, Application US/10712479
; Publication No. US20040209283A1
; GENERAL INFORMATION:
; APPLICANT: Yagi, Yoshihiko
; APPLICANT: Sheets, Michael P.
; APPLICANT: Wells, Peter A.
; APPLICANT: Shelly, John A.
; APPLICANT: Poorman, Roger A.
; APPLICANT: Epps, Dennis E.
; TITLE OF INVENTION: A Continuous-Read Assay for the Detection of De Novo HCV RNA Polyn
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 02-333-A
; CURRENT APPLICATION NUMBER: US/10/712,479
; CURRENT FILING DATE: 2003-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: C-terminally truncated HCV NS5B polymerase (C delta 21 NS5B)
US-10-712-479-4

Query Match   100.0%;   Score 3055;   DB 16;   Length 580;
Best Local Similarity   100.0%;   Pred. No. 4.1e-269;
Matches 580;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHNLVYSTTSRSASLRQKKVTFDR 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MASMSMSYTWTGALITPCAAEESKLPINALSNSLLRHNLVYSTTSRSASLRQKKVTFDR 60

```
Qy    61  LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  LQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAV  120

Qy   121  NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  NHIRSVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY  180

Qy   181  DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  DVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVE  240

Qy   241  ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  ESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCY  300

Qy   301  LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQP  360

Qy   361  EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  EYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYA  420

Qy   421  PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  PTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHS  480

Qy   481  YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  YSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKL  540

Qy   541  TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  580
          |||||||||||||||||||||||||||||||||||||||
Db   541  TPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  580
```

RESULT 2
US-10-842-046-1
; Sequence 1, Application US/10842046
; Publication No. US2005003348A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Hepatits C virus NS5B Polymerase Inhibitor Binding Pocket
; FILE REFERENCE: 13/123
; CURRENT APPLICATION NUMBER: US/10/842,046
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/469,604
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-842-046-1

```
Query Match             96.8%;  Score 2958;  DB 17;  Length 576;
Best Local Similarity   96.0%;  Pred. No. 2.8e-260;
Matches  553;  Conservative  16;  Mismatches  7;  Indels  0;  Gaps  0;

Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  60

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          |||||||||||||||||||||||||:|||||||||||||||||||||||:||||||||
Db    61  DDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSRAVNHIR  120

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||:||||||:||::|||||:|||||||||||||||||||||||||||||||||||||||
Db   121  SVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180
```

```
Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          |||||||||||||||||||||||||||||:|||||||||||||||||:||||||||
Db   181  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  240

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          |||||||||||||||||||||:|||::||||||||||||||||||||||||||||||:
Db   241  QCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAT  300

Qy   305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
          |||||||||||||||||||||||||||||||||||:||:||||||||||||||||||||
Db   301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPPGDPPQPEYDL  360

Qy   365  ELITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
          |||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db   361  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPINSWLGNIIMYAPTLW              420

Qy   425  ARMILMTHFFSILLAGTQEDQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
          |||||||||||||:::||||||||||||||||||||||||||||||||||||||||:||||||
Db   421  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFTLHSYSPG     480

Qy   485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
          ||||||||||||||||||:|||||||||||||||||||||:||:|||||||||||||||
Db   481  EINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIP  540

Qy   545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  580
          |||:|||||||||||||||||||||||||||||||
Db   541  AASQLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH  576
```

RESULT 3
US-10-211-455-1
; Sequence 1, Application US/10211455
; Publication No. US2003010862A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: DIRECT BINDING ASSAY FOR IDENTIFYING
; TITLE OF INVENTION: INHIBITORS OF HCV POLYMERASE
; FILE REFERENCE: 13/088
; CURRENT APPLICATION NUMBER: US/10/211,455
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/310,272
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 578
; TYPE: PRT
; ORGANISM: HCV
US-10-211-455-1

```
Query Match             96.1%;  Score 2936;  DB 14;  Length 578;
Best Local Similarity   95.5%;  Pred. No. 2.9e-258;
Matches  552;  Conservative  18;  Mismatches  6;  Indels  2;  Gaps  1;

Qy     5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
          |||||||||||||||||||:|||||||||||:||:||||||||||||:||||||||||||
Db     1  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMYSTTSRSAALRQKKVTFDRLQVL   60

Qy    65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
          ||||||||||||||||||||||||||||||||||||||||||||||||:|||:||:||||
Db    61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  120

Qy   125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy   185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
          |||||||||||||||||||||||||||||:|||||||||||||||||:||||||||
Db   181  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  240

Qy   245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
          |||||||||||||:|||||||:|||::||||||||||||||||||||||||||||||:
Db   241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300
```

RESULT 4
US-10-471-164-3
; Sequence 3, Application US/10471164
; Publication No. US20040110126A1
; GENERAL INFORMATION:
; APPLICANT: George KUKOLJ and Ginette MCKERCHER
; TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
; FILE REFERENCE: 13/094
; CURRENT APPLICATION NUMBER: US/10/471,164
; CURRENT FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/274,374
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 578
; TYPE: PRT
; ORGANISM: HCV
US-10-471-164-3

Query Match   96.1%;   Score 2936;   DB 16;   Length 578;
Best Local Similarity   95.5%;   Pred. No. 2.9e-258;
Matches 552;   Conservative 18;   Mismatches 6;   Indels 2;   Gaps 1;

```
Qy    5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
Db    1  SMSYTWTGALITPCAAEESQLPINALSNSLLRHHNMVYSTTSRSAALRQKKVTFDRLQVL  60

Qy   65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR  124
Db   61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR  120

Qy  125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  184
Db  121  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS  180

Qy  185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY  244
Db  181  TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY  240

Qy  245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA  304
Db  241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS  300

Qy  305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL  364
Db  301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL  360

Qy  365  ELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLW  424
Db  361  ELITCSSSNVSVAHDASGKRVYYLTRDPTTPINSWLGNIIMYAPTLW  420

Qy  425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  484
Db  421  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG  480

Qy  485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP  544
Db  481  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIP  540

Qy  545  AASRLDLSGWFVAGYSGGDIYHSLSRARPR--HHHHHH  580
Db  541  AASRLDLSGWFVAGYNGGDIYHSLSRARPRLEHHHHHH  578
```

RESULT 5
US-10-170-131-1
; Sequence 1, Application US/10170131
; Publication No. US20030171874A1
; GENERAL INFORMATION:
; APPLICANT: Lesburg, Charles A.
; APPLICANT: Cable, Michael
; APPLICANT: Hong, Zhi
; APPLICANT: Mannarino, Anthony
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Compositions of Hepatitis C Virus NS5B Polymerase and
; TITLE OF INVENTION: Methods for Crystallizing Same
; FILE REFERENCE: IN0988 US
; CURRENT APPLICATION NUMBER: US/10/170,131
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (335)
; OTHER INFORMATION: amino acids at 335, 344 and 550 may be threonine
; OTHER INFORMATION: or valine at 335, valine or alanine at 344 and
; OTHER INFORMATION: arginine or glutamine at 550
US-10-170-131-1

Query Match   94.2%;   Score 2879;   DB 14;   Length 576;
Best Local Similarity   96.0%;   Pred. No. 4.5e-253;
Matches 545;   Conservative 11;   Mismatches 12;   Indels 0;   Gaps 0;

```
Qy    7  SYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVLDD  66
Db    9  SYTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDD  68

Qy   67  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSV  126
Db   69  HYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHIHSV  128

Qy  127  WKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  186
Db  129  WKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL  188

Qy  187  PQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQC  246
Db  189  PQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRVEESIYQC  248

Qy  247  CDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAA  306
Db  249  CDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAA  308

Qy  307  CRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLEL  366
Db  309  CRAAKLQDCTMLVNGDDLVVICESAGKQEDAASLRXFTEAMTRYSAPPGDPPQPEYDLEL  368

Qy  367  ITSCSSNVSVAHDASGKRVYYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  426
```

```
Db    369  ITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR        428
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKL------      534

Qy    427  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI      486
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          580
Db    429  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI      488
               |:|:|
Db    535  -AAALE--------------HHHHHH                                    545

Qy    487  NRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAA      546

RESULT 7
US-09-921-397-107
; Sequence 107, Application US/09921397
; Patent No. US2002015148A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
;  TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
;  TITLE OF INVENTION: applications thereof
;  FILE REFERENCE: B4809A - JAZ
;  CURRENT APPLICATION NUMBER: US/09/921,397
;  CURRENT FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: EP 00402225.7
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 156
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 107
;  LENGTH: 300
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-09-921-397-107

Query Match            45.5%;  Score 1391;  DB 9;   Length 300;
Best Local Similarity  87.7%;  Pred. No. 7.4e-118;
Matches  263;  Conservative  13;  Mismatches  24;  Indels  0;  Gaps  0;

Qy     84  AKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSVWKDLLEDTDTPIQTTIM     143
               ||||||||||||||||||||||||||||||||||| :|:||||||:|:|||| ||:|||
Db      1  ANLLSVEEACSLTPPHSAKSKFGYGAKDVRSLSSRAVAHINSVWKDLLEDSVTPIDTTIM      60

Qy    144  AKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQ     203
               |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db     61  AKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSKIPLAVMGSSYGFQYSPGQ     120

Qy    204  RVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPEAROAIRSLTER     263
               |||| |:|||||||||||||||||||||||:|||||||:||||| | |:|||||||||
Db    121  RVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTER     180

Qy    264  LYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDD     323
               ||||||:|||:|||||||||||||||||||||||||||||:||||:||| ::::|:||
Db    181  LYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDD     240

Qy    324  LVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGK     383
               ||||||||||||||| ||||||:|||||||||||||||||||||||||||::|||||||
Db    241  LVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPPQPEYDLELITSCSSNVSVAHDGAGK     300

RESULT 8
US-10-927-520-17
; Sequence 17, Application US/10927520
; Publication No. US2005006970A1
; GENERAL INFORMATION:
;  APPLICANT: Innogenetics N.V.
;  TITLE OF INVENTION: New HCV clade and prototype sequences thereof
;  FILE REFERENCE: 157
;  CURRENT APPLICATION NUMBER: US/10/927,520
;  CURRENT FILING DATE: 2004-08-27
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;  LENGTH: 352
;  TYPE: PRT
;  ORGANISM: hepatitis C virus
US-10-927-520-17

Query Match            42.0%;  Score 1283;  DB 17;  Length 352;
```

```
Db    481  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKL------      534

Qy    545  AASRLDLSGWFVAGYSGGDIYHSLSRARPRHHHHHH                          580

Db    535  -AAALE--------------HHHHHH                                    545
```

```
Db    369  ITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR        428

Qy    427  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI      486
Db    429  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEI      488

Qy    487  NRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIPAA      546
Db    489  NRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAA      548

Qy    547  SRLDLSGWFVAGYSGGDIYHSLSRARPR      574
Db    549  SXLDLSGWFVAGYSGGDIYHSLSRARPR      576

RESULT 6
US-10-471-164-4
; Sequence 4, Application US/10471164
; Publication No. US2004110126A1
; GENERAL INFORMATION:
;  APPLICANT: George KUKOLJ and Ginette MCKERCHER
;  TITLE OF INVENTION: HCV POLYMERASE INHIBITOR ASSAY
;  FILE REFERENCE: 13/094
;  CURRENT APPLICATION NUMBER: US/10/471,164
;  CURRENT FILING DATE: 2003-09-05
;  PRIOR APPLICATION NUMBER: 60/274,374
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;  LENGTH: 545
;  TYPE: PRT
;  ORGANISM: HCV
US-10-471-164-4

Query Match            89.5%;  Score 2734.5;  DB 16;  Length 545;
Best Local Similarity  90.1%;  Pred. No. 6.2e-240;
Matches  519;  Conservative  19;  Mismatches  7;  Indels  31;  Gaps  2;

Qy      5  SMSYTWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL      64
               |||||||||||||||||||||:||||||||||||||  :||||||||||||||||||||
Db      1  SMSYTWTGALITPCAAEESQLPINALSNSLVRHRNMVYSTTSRSAALRQKKVTFDRLQVL      60

Qy     65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR     124
               ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||:
Db     61  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVDHIR     120

Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS     184
               |||||||||:|:||:||||||||||||||||||||||||||||||||||||||||||||
Db    121  SVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS     180

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY     244
               |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db    181  TLPQAVMGSSYGFQYSPKQRVEFLVNAWKSKKCPMGFSYDTRCFDSTVTESDIRVEESIY     240

Qy    245  QCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA     304
               |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||:
Db    241  QCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAS     300

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDL     364
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db    301  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAANLRVFTEAMTRYSAPPGDLPQPEYDL     360

Qy    365  ELITSCSSNVSVAHDASGKRVYLTRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLW      424
               ||||||||||||||||||||||||||||:||||||||||||||:|||||||:||||||
Db    361  ELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLW      420

Qy    425  ARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG     484
               ||:||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  ARMVLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPG     480

Qy    485  EINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAAICGKYLFNWAVRTKLKLTPIP     544
```

Best Local Similarity 71.5%; Pred. No. 6.5e-108; Indels 0; Gaps 0;
Matches 243; Conservative 38; Mismatches 59;

```
Qy      5  SMSYTWTGALITPCAABESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQVL  64
              ||||||||| ||:||  |||||||:||  |:|||||:|:|::||
Db     13  SMSYSWTGALVTPSGPEEERLPINALSNTMLRHYNMVYSTTSRSAQARAKKVTFDRLQVL  72

Qy     65  DDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIR 124
              |||| :| ::||| |||||||:|:||| ||||:|:: ||
Db     73  DDHYKRTLDDVKAKAGVTARLLTLEEAAALTPTHSARSKFGYGAKDVRALAPKAVTDIK 132

Qy    125  SVWKDLLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 184
              |||: ||  ||:|||||||||||||||||||||||||||||||||||||||||
Db    133  GVWKNLLTDKTTPIPTSIMAKNEVFCVNPAKGGKKRPARLIYYPDLGVRVCEKRALYDLAQ 192

Qy    185  TLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIY 244
              :|||||||||||||||||||||||||||||||||||||||| ::| |||||||
Db    193  KLPQAVMGSAYGFQYSPAQRVDLLVKTWESKRTPMGFSYDTRCFDSTVTEQDIRTETAIY 252

Qy    245  QCCDLLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAA 304
              ||| | ||||:|||:|||||||||||||||||||||||:|||||||||||||||
Db    253  QCCNLDPEARTAIDALTERLYVGGPMFNSKGEKVGYRRCRASGVFPTSMGNTMTCYIKAK 312

Qy    305  AACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFT 344
              |||| || |:||||||||||||||||||| :|||||
Db    313  AASAAAGLSGADFLVCGDDLVVICESKGVDQDRAALRAFT 352
```

RESULT 9
US-09-921-397-109
; Sequence 109, Application US/09921397
; Patent No. US2002015148A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
; TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
; TITLE OF INVENTION: applications thereof
; FILE REFERENCE: B4809A - JAZ
; CURRENT APPLICATION NUMBER: US/09/921,397
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP 00402225.7
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-921-397-109

Query Match 40.3%; Score 1232; DB 9; Length 260;
Best Local Similarity 89.2%; Pred. No. 1.9e-103;
Matches 232; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

```
Qy    130  LLEDTDTPIQTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVS 189
              ||||: |||||||||||||||||||||||||||||||||||||||||||||||
Db      1  LLEDSVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSKLPLA  60

Qy    190  VMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDL 249
              |||||||||||||||||||||||||||||||||||||||||| ||||||:|||||
Db     61  VMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIYQCCDL 120

Qy    250  APEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAACRA 309
              ||:||:|||:||||||||||| ||:||:||||||||||||:|||||||||:|||| ||
Db    121  DPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLRVFTEAMTRYSAPPGDPPQPEYDLELITS 180

Qy    310  AKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITS 369
              | ||||||:|||||||||||||||:||||:|||||||||||||||||||||||||||||||
Db    181  AGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITS 240

Qy    370  CSSNVSAHDASGKRVYYLT 389
              |||||:||| ||||||||
Db    241  CSSNVSVAHDGAGKRVYYLT 260
```

RESULT 10
US-10-211-455-3
; Sequence 3, Application US/10211455
; Publication No. US20030108862A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim (Canada) Ltd.
; TITLE OF INVENTION: DIRECT BINDING ASSAY FOR IDENTIFYING
; TITLE OF INVENTION: INHIBITORS OF HCV POLYMERASE
; FILE REFERENCE: 13/088
; CURRENT APPLICATION NUMBER: US/10/211,455
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/310,272
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 576
; TYPE: PRT
; ORGANISM: HCV
US-10-211-455-3

Query Match 31.9%; Score 975.5; DB 14; Length 576;
Best Local Similarity 38.7%; Pred. No. 1.4e-79;
Matches 230; Conservative 93; Mismatches 235; Indels 35; Gaps 12;

```
Qy      4  MSMSYTWTGALITPCAABESKLPINALSNSLLRHHNLVYSTTSRSASLRQKKVTFDRLQV  63
              ||||||||  :||:| | |||   |:::   |||:||||||||||||:||||:||
Db      1  MSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPL  59

Qy     64  LDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKF-GYGAKDVRSLSSR-AVN 121
              |||||   |:||:|:|  :|  ||:  | |||||:|:  :|| ||||||||:||| |:
Db     60  FPPSYHKQVRLAKEKASKVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVL 119

Qy    122  HIRSVWKDLLEDTDTPI---QTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMA 178
              :: :   :|:  | ||  |||||:||:|| | || | :|:|||:||| || || |:||
Db    120  DIQK----CVEAGBIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMY 175

Qy    179  LYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKAKKCPMGFSYDTRCFDSTVTENDIR 238
              |||:    |:||  | :   |  |::|::|:   |:  |  ||:|     :|| :|:
Db    176  YGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTITPEDIM 232

Qy    239  VEESIYQCCDLLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVLTTSCGNTLT 298
              |  :|| |  ||   |:|||:| :||  |    |:    |||:  |::| || | |:|:
Db    233  VETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLT 292

Qy    299  CYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPP 358
              | |::|   |  ||: | ::|  |||| |:|  | :|:|| |:| |:::|    || | |
Db    293  CWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVP 352

Qy    359  QPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTVPVNSWLGNIIM 418
              ||::| |||  :  |   || :|| :|    ||  | |:|  |: |:
Db    353  QPKYSLEBLTSCCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIH 412

Qy    419  YAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSL 478
              |||    |:  |:| |     |  | :|  | || :| | ||  ||::||:
Db    413  HYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEARFSV 472

Qy    479  HSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGRAAICGKYLFNWAVRTKL 538
              |   |  ||:|  || | :| |:|   |:  ::| |    |  |  ::
Db    473  VRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR- 530

Qy    539  KLTPIPAASRLDLSGWFVAGYS--------GGDIYHSLSR--ARPRHHHHHH 580
              |:|        ||  |:|         |  ::   |  | |||||||
Db    531  ---PLP-----DLDKTSVARYTTFNYCDVYSPEGDVFTPQRRLQKLEHHHHHH 576
```

RESULT 11
US-09-921-397-108
; Sequence 108, Application US/09921397
; Patent No. US2002015148A1
; GENERAL INFORMATION:

```
;  APPLICANT: HYBRIGENICS
;  TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
;  TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
;  TITLE OF INVENTION: applications thereof
;  FILE REFERENCE: B4809A - JAZ
;  CURRENT APPLICATION NUMBER: US/09/921,397
;  CURRENT FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: EP 00402225.7
;  PRIOR FILING DATE: 2000-08-03
;  SOFTWARE: PatentIn Ver. 2.1
;  NUMBER OF SEQ ID NOS: 156
;  SEQ ID NO 108
;  LENGTH: 199
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-09-921-397-108

Query Match    30.0%;  Score 915;  DB 9;  Length 199;
Best Local Similarity  85.9%;  Pred. No. 9.9e-75;
Matches 171;  Conservative 11;  Mismatches 17;  Indels 0;  Gaps 0;

Qy  90 EEACKLTPPHSAKSKFGYGAKDVRSLSSRAVNHIRSVWKDLLEDTDTPIQTTIMAKQNEVF 149
        ||||:|||||||||||||||||||| :||:||||||:||:||||||||||||||||||
Db   1 EEACSLTPPHSAKSKFGYGAKDVRCHARKAVAHINSVWKDLLEDSVTPIDTTIMAKQNEVF  60

Qy 150 CVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLIV 209
        ||||||||||||||||||||||||||||||||||||:||:||:|||||||||:|||||||
Db  61 CVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEFLV  120

Qy 210 NTWKAKKCPMGFSYDTRCFDSTVTENDIRVBESIYQCCDLAPEARQAIRSLTERLYVGGP  269
        :|||:|: :|||||||||||||||:|:|||||||||:||:|||:||||||||||
Db 121 QAWKSKKTPMGFSYDTRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGP  180

Qy 270 MTNSKGQNCGYRRCRASGV 288
        :|||:|:|||||||||||
Db 181 LTNSRGENCGYRRCRASGV 199
```

```
RESULT 12
US-09-921-397-110
;  Sequence 110, Application US/09921397
;  Patent No. US2002151484A1
;  GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
;  TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
;  TITLE OF INVENTION: applications thereof
;  FILE REFERENCE: B4809A - JAZ
;  CURRENT APPLICATION NUMBER: US/09/921,397
;  CURRENT FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: EP 00402225.7
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 156
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 110
;  LENGTH: 127
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-09-921-397-110

Query Match    20.4%;  Score 623;  DB 9;  Length 127;
Best Local Similarity  89.8%;  Pred. No. 2.1e-48;
Matches 114;  Conservative 9;  Mismatches 4;  Indels 0;  Gaps 0;

Qy 326 VICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCCSSNVSVAHDASGKRV 385
        |||||||:||||||:|||||||||||||||||||||||||||||||||||||||:|||:
Db   1 VICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPEYDLELITSCCSSNVSVAHDGAGKRV  60

Qy 386 YLITRDPTVPLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMWTHFFSILLAQBQLEK 445
        ||||||:|||||||||||||||||||||||::|||||||||||||||:||:::||:|
Db  61 YLITRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLWARMILMWTHFFSVLIARDQLEQ 120

Qy 446 ALDCQIY 452
```

```
Db 121 ALNCEIY 127
```

```
RESULT 13
US-09-921-397-106
;  Sequence 106, Application US/09921397
;  Patent No. US20020151484A1
;  GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  TITLE OF INVENTION: SID nucleic acids and polypeptides selected from a
;  TITLE OF INVENTION: pathogenic strain of the hepatitis C virus and
;  TITLE OF INVENTION: applications thereof
;  FILE REFERENCE: B4809A - JAZ
;  CURRENT APPLICATION NUMBER: US/09/921,397
;  CURRENT FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: EP 00402225.7
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 156
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 106
;  LENGTH: 137
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-09-921-397-106

Query Match    19.1%;  Score 585;  DB 9;  Length 137;
Best Local Similarity  83.9%;  Pred. No. 6.9e-45;
Matches 115;  Conservative 7;  Mismatches 15;  Indels 0;  Gaps 0;

Qy   9 TWTGALITPCAAEESKLPINALSNSLLRHHNLVYSTTSRSASLRQKVTFDRLQVLDDHY  68
        :||||:|||||||:||||||||||||||||||||||||||: | ||:|||||||||:||
Db   1 SWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVLDSHY  60

Qy  69 RDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLLSSRAVNHIRSVWK 128
        :|||:|:||| |:|||:|||||||:||||||||||||||||||:| || |:|:||:||||
Db  61 QDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVAHINSVWK  120

Qy 129 DLLEDTDTPIQTTIMAK 145
        ||||| :||:|||||||
Db 121 DLLEDSVTPIDTTIMAK 137
```

```
RESULT 14
US-09-294-121A-82
;  Sequence 82, Application US/09294121A
;  Patent No. US2002006942A1
;  GENERAL INFORMATION:
;  APPLICANT: MAERTENS, GEERT; STUYVER, LIEVEN;
;  APPLICANT: ROSSAU, RUDI; VAN HEUVERSWYN, HUGO
;  TITLE OF INVENTION: PROCESS FOR TYPING OF HCV
;  TITLE OF INVENTION: ISOLATES
;  NUMBER OF SEQUENCES: 97
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: BIERMAN & MUSERLIAN
;  STREET:  600 THIRD AVENUE
;  CITY:  NEW YORK
;  STATE:  NEW YORK
;  COUNTRY: USA
;  ZIP: 10016
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE: ASCII
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/294,121A
;  FILING DATE:
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 08/256,568
;  FILING DATE:  18-JUL-1994
;  APPLICATION NUMBER: PCT/EP93/03325
```

RESULT 15
US-09-899-082A-82
; Sequence 82, Application US/09899082A
; Patent No. US2002010663BA1
; GENERAL INFORMATION:
;   APPLICANT: MAERTENS, GEERT; STUYVER, LIEVEN;
;   ROSSAU, RUDI; VAN HEUVERSWYN, HUGO
;   TITLE OF INVENTION: PROCESS FOR TYPING OF HCV
;   ISOLATES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BIERMAN & MUSERLIAN
;   STREET: 600 THIRD AVENUE
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: USA
;   ZIP: 10016
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/899,082A
;   FILING DATE: 06-Jul-2001
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/378,900
;   FILING DATE: <Unknown>
;   APPLICATION NUMBER: 08/256,568
;   FILING DATE: 18-JUL-1994
;   APPLICATION NUMBER: PCT/EP93/03325
;   FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: EP/93/402,129.6
;   FILING DATE: 31-AUG-1993
;   APPLICATION NUMBER: EP/92/403,222.0
;   FILING DATE: 27-NOV-1992
; ATTORNEY/AGENT INFORMATION:
;   NAME: CHARLES A. MUSERLIAN
;   REGISTRATION NUMBER: 19,683
;   REFERENCE/DOCKET NUMBER: 410.004
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 661-8000
;   TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 82:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 113 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   IMMEDIATE SOURCE:
;   CLONE: be90
;   POSITION IN GENOME:
;   MAP POSITION: ns5
US-09-294-121A-82

Query Match   18.7%;   Score 570; DB 9; Length 113;
Best Local Similarity   95.6%; Pred. No. 1.2e-43;
Matches 108;   Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
        ||||||||||||||||||||||| ||:||:|||||||||:|||||||||||||||||||
Db   1 STVTENDIRVEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy 290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
        |||||||||||| ||||:|||||||||||| ||||||||||||||||||||||
Db   61 TTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRV 113

;   FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: EP/93/402,129.6
;   FILING DATE: 31-AUG-1993
;   APPLICATION NUMBER: EP/92/403,222.0
;   FILING DATE: 27-NOV-1992
; ATTORNEY/AGENT INFORMATION:
;   NAME: CHARLES A. MUSERLIAN
;   REGISTRATION NUMBER: 19,683
;   REFERENCE/DOCKET NUMBER: 410.004
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 661-8000
;   TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 82:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 113 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   IMMEDIATE SOURCE:
;   CLONE: be90
;   POSITION IN GENOME:
;   MAP POSITION: ns5
;   SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-899-082A-82

Query Match   18.7%;   Score 570; DB 9; Length 113;
Best Local Similarity   95.6%; Pred. No. 1.2e-43;
Matches 108;   Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 230 STVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYVGGPMTNSKGQNCGYRRCRASGVL 289
        ||||||||||||||||||||||| ||:||:|||||||||:|||||||||||||||||||
Db   1 STVTENDIRVEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVL 60

Qy 290 TTSCGNTLTCYLKAAAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRV 342
        |||||||||||| ||||:|||||||||||| ||||||||||||||||||||||
Db   61 TTSCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAASLRV 113

Search completed: September 22, 2005, 14:58:37
Job time : 175 secs

This Page Blank (uspto)